



US009441019B2

(12) **United States Patent**
Nabel et al.(10) **Patent No.:** **US 9,441,019 B2**
(45) **Date of Patent:** **Sep. 13, 2016**(54) **INFLUENZA HEMAGGLUTININ
PROTEIN-BASED VACCINES**(71) Applicant: **THE UNITED STATES OF
AMERICA, as represented by THE
SECRETARY, DEPARTMENT OF
HEALTH & HUMAN SERVICES,**
Washington, DC (US)(72) Inventors: **Gary J. Nabel**, Washington, DC (US);
Masaro Kanekiyo, Silver Spring, MD
(US); **Chih-Jen Wei**, Potomac, MD
(US); **Patrick M. McTamney**,
Bethesda, MD (US); **Hadi M. Yassine**,
Rockville, MD (US); **Jeffrey C.
Boyston**, Clarksburg, MD (US)(73) Assignee: **The United States of America, as
represented by the Secretary,
Department of Health and Human
Services**, Washington, DC (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **14/346,849**(22) PCT Filed: **Sep. 24, 2012**(86) PCT No.: **PCT/US2012/056822**

§ 371 (c)(1),

(2) Date: **Mar. 24, 2014**(87) PCT Pub. No.: **WO2013/044203**PCT Pub. Date: **Mar. 28, 2013**(65) **Prior Publication Data**

US 2014/0302079 A1 Oct. 9, 2014

Related U.S. Application Data

(60) Provisional application No. 61/538,663, filed on Sep. 23, 2011, provisional application No. 61/661,209, filed on Jun. 18, 2012.

(51) **Int. Cl.****C07K 14/205** (2006.01)**A61K 39/12** (2006.01)**A61K 39/145** (2006.01)**A61K 39/00** (2006.01)**G01N 33/569** (2006.01)**C07K 14/005** (2006.01)**A61K 39/02** (2006.01)(52) **U.S. Cl.**CPC **C07K 14/205** (2013.01); **A61K 39/12**(2013.01); **A61K 39/145** (2013.01); **C07K****14/005** (2013.01); **G01N 33/56983** (2013.01);**A61K 39/105** (2013.01); **A61K 2039/55566**(2013.01); **A61K 2039/6031** (2013.01); **A61K****2039/70** (2013.01); **C07K 2319/00** (2013.01);**C12N 2760/16034** (2013.01); **C12N 2760/16134** (2013.01); **C12N 2760/16234**

(2013.01)

(58) **Field of Classification Search**

None

See application file for complete search history.

(56)

References Cited

U.S. PATENT DOCUMENTS

7,094,598 B2	8/2006	Nabel et al.
7,097,841 B2	8/2006	Carter et al.
7,608,268 B2	10/2009	Carter et al.
2002/0054882 A1	5/2002	Okuno et al.
2005/0042229 A1	2/2005	Yang et al.
2005/0108791 A1	5/2005	Edgerton
2006/0251679 A1*	11/2006	Carter et al. 424/204.1
2007/0082054 A1	4/2007	Van Den Mooter et al.
2007/0224205 A1	9/2007	Powell et al.
2008/0299151 A1	12/2008	Fomsgaard
2009/0233377 A1	9/2009	Iwahori et al.
2010/0137412 A1	6/2010	Zhou et al.
2010/0285982 A1	11/2010	Golding et al.
2010/0297174 A1	11/2010	Garcia-Sastre et al.
2011/0038025 A1	2/2011	Naitou et al.
2011/0059130 A1	3/2011	Yusibov
2011/0177122 A1	7/2011	Nabel et al.
2011/0212128 A1	9/2011	Galarza et al.

FOREIGN PATENT DOCUMENTS

EP	1504037	12/2009
WO	WO 03/094849	11/2003

(Continued)

OTHER PUBLICATIONS

Cohen et al., *Neoplasia*, Feb. 2005, 7(2):109-117.*
A3KF33, UniProtKB A3KF33_157A5, Sep. 21, 2011 [online].
[Retrieved on Feb. 26, 2013]. Retrieved from the internet <URL:
<http://www.uniprot.org/uniprot/A3KF33.txt?version=36>>.
Bachmann, M.F., et al., "Neutralizing antiviral B cell responses," *Annu Rev Immunol*, 1997, 15:235-270.
Caton, A.J., et al., "The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype)," *Cell*, 1982, 31:417-427.

(Continued)

Primary Examiner — Nicole Kinsey White(74) *Attorney, Agent, or Firm* — Sheridan Ross P.C.(57) **ABSTRACT**

Novel vaccines are provided that elicit broadly neutralizing anti-influenza antibodies. Some vaccines comprise nanoparticles that display hemagglutinin trimers from influenza virus on their surface. The nanoparticles comprise fusion proteins comprising a monomeric subunit of ferritin joined to at least a portion of an influenza hemagglutinin protein. Some portions comprise the ectodomain while some portions are limited to the stem region. The fusion proteins self-assemble to form the hemagglutinin-displaying nanoparticles. Some vaccines comprise only the stem region of an influenza hemagglutinin protein joined to a trimerization domain. Such vaccines can be used to vaccinate an individual against infection by heterologous influenza viruses and influenza virus that are antigenically divergent from the virus from which the nanoparticle hemagglutinin protein was obtained. Also provided are fusion proteins and nucleic acid molecules encoding such proteins.

(56)

References Cited

FOREIGN PATENT DOCUMENTS

WO WO 2011/035422 * 3/2011
WO WO 2011/044152 A1 4/2011

OTHER PUBLICATIONS

- COLT38, UniProtKB COLT38_9INFB, Sep. 21, 2011 [online]. [Retrieved on Feb. 26, 2013]. Retrieved from the internet <URL: <http://www.uniprot.org/uniprot/COLT38.txt?version=18>>.
- Corti, D., et al., "Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine," *J Clin Invest*, 2010, 120:1663-1673.
- Corti, D., et al., "A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins," *Science*, 2011, 333:850-856.
- Dintzis, H.M. et al., "Molecular determinants of immunogenicity: the immunon model of immune response," *Proc Natl Acad Sci USA*, 1976, 73:3671-3675.
- Ekiert, D.C., et al., "A highly conserved neutralizing epitope on group 2 influenza A viruses," *Science*, 2011, 333:843-850.
- Ekiert, D.C., et al., "Antibody recognition of a highly conserved influenza virus epitope," *Science*, 2009, 324:246-251.
- Haynes, J.R., "Influenza virus-like particle vaccines," *Expert Rev Vaccines*, 2009, 8:435-445.
- Kashyap, A.K., et al., "Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies," *Proc Natl Acad Sci USA*, 2008, 105:5986-5991.
- Kong, W.P., et al., "Protective immunity to lethal challenge of the 1918 pandemic influenza virus by vaccination," *Proc Natl Acad Sci USA*, 2006, 103:15987-15991.
- Krause, J.C., et al., "A broadly neutralizing human monoclonal antibody that recognizes a conserved, novel epitope on the globular head of the influenza H1N1 virus hemagglutinin," *J Virol*, 2011, 85:10905-10908.
- Lambert, L.C., et al., "Influenza vaccines for the future," *N Engl J Med*, 2010, 363, 2036-2044.
- Li, C.Q. et al., "Ferritin nanoparticle technology: A new platform for antigen presentation and vaccine development," *Industrial Biotechnol* 2, 143-147 (2006).
- Meldrum, F.C., et al., "Magnetoferitin: in vitro synthesis of a novel magnetic protein," *Science*, 1992, 257:522-523.
- Nabel, G.J., et al., "Induction of unnatural immunity: prospects for a broadly protective universal influenza vaccine," *Nat Med*, 2010, 16:1389-1391.
- Okuno, Y., et al., "A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains," *J Virol*, 1993, 67:2552-2558.
- Roldao, A., et al., "Virus-like particles in vaccine development," *Expert Rev Vaccines*, 2010, 9:1149-1176.
- Sheridan, C., "Flu vaccine makers upgrade technology—and pray for time," *Nat Biotechnol*, 2009, 27:489-491.
- Steel, J. et al., "Influenza Virus Vaccine Based on the Conserved Hemagglutinin Stalk Domain," *MBio* 1, e0018 (2010).
- Sui, J., et al., "Structural and functional bases for broad-spectrum neutralization of avian and human influenza a viruses," *Nat Struct Mol Biol*, 2009, 16:265-273.
- Treanor, J.J., et al., "Safety and immunogenicity of a recombinant hemagglutinin vaccine for H5 influenza in humans," *Vaccine*, 2001, 19:1732-1737.
- Treanor, J.J., "Safety and immunogenicity of a baculovirus-expressed hemagglutinin influenza vaccine: a randomized controlled trial," *JAMA*, 2007, 297:1577-1582.
- Wang, T.T., et al., "Broadly protective monoclonal antibodies against H3 influenza viruses following sequential immunization with different hemagglutinins," *PLoS Pathog*, 2010, vol. 6, Issue 2, e1000796.
- Wei, C.J., et al., "Induction of broadly neutralizing H1N1 influenza antibodies by vaccination," *Science*, 2010, 329:1060-1064.
- Wei, C.J., et al., "Comparative efficacy of neutralizing antibodies elicited by recombinant hemagglutinin proteins from avian H5N1 influenza virus," *J Virol*, 2008, 82:6200-6208.
- Wei, C.J., et al., "Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design," *Sci Transl Med*, 2010, 2, 24ra21.
- Whittle, J.R., et al., "Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin," *Proc Natl Acad Sci USA*, 2011, 108:14216-14221.
- WHO Reference on Animal Influenza Diagnosis and Surveillance, 2002, Department of Communicable Disease Surveillance and Response, World Health Organization).
- Wu, C.Y., et al., "Mammalian expression of virus-like particles for advanced mimicry of authentic influenza virus," *PLoS One* 5, 2010, e9784.
- Xiong, A.S., et al., "PCR-based accurate synthesis of long DNA sequences," *Nat Protoc*, 2006, 1(2):791-797.
- Yamashita, I., et al., "Ferritin in the field of nanodevices," *Biochim Biophys Acta*, 2010, 1800:846-857.
- Yang, Z.Y., et al., "Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity," *Science*, 2007, 317:825-828.
- Zhang, Y., et al., "Self-Assembly in the Ferritin Nano-Cage Protein Superfamily," *Int. J. Mol. Sci.*, 2011, 12:5406-5421.
- Invitation to remedy deficiencies pursuant to Rule 30(3) EPC / Rules 163(3) EPC, dated May 23, 2014, in European Patent Application No. 12834398.5.

* cited by examiner

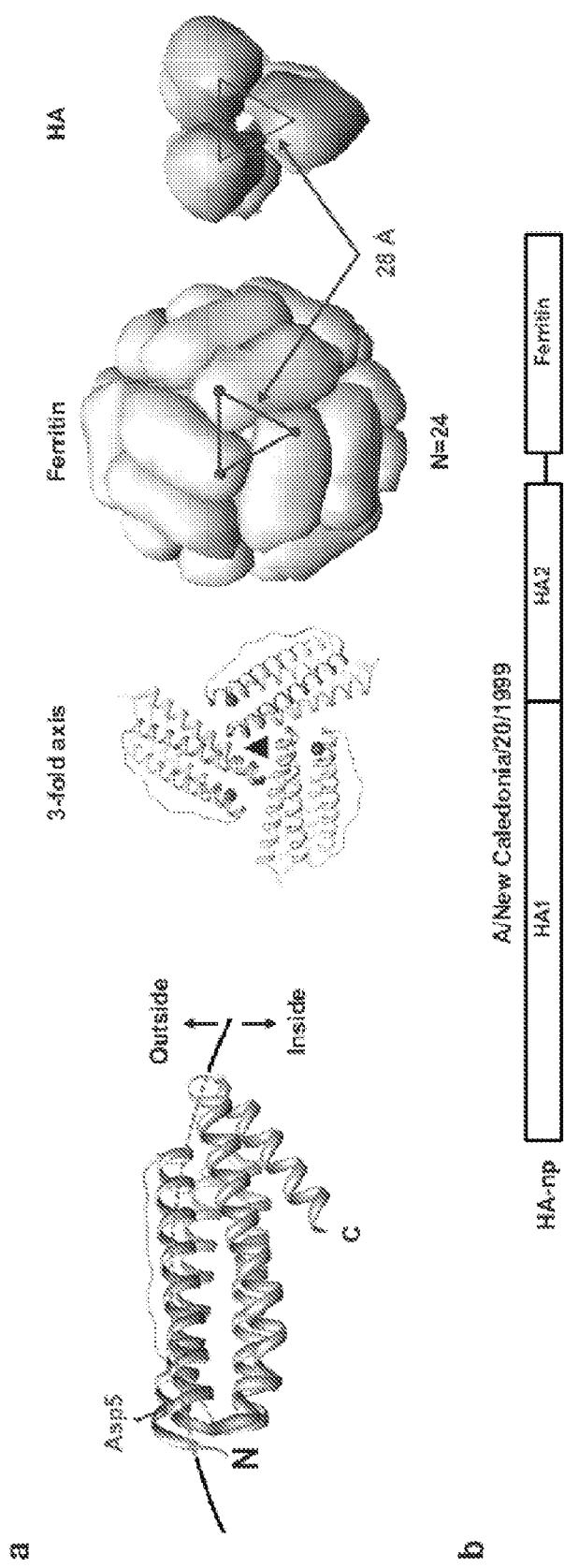


Fig. 1

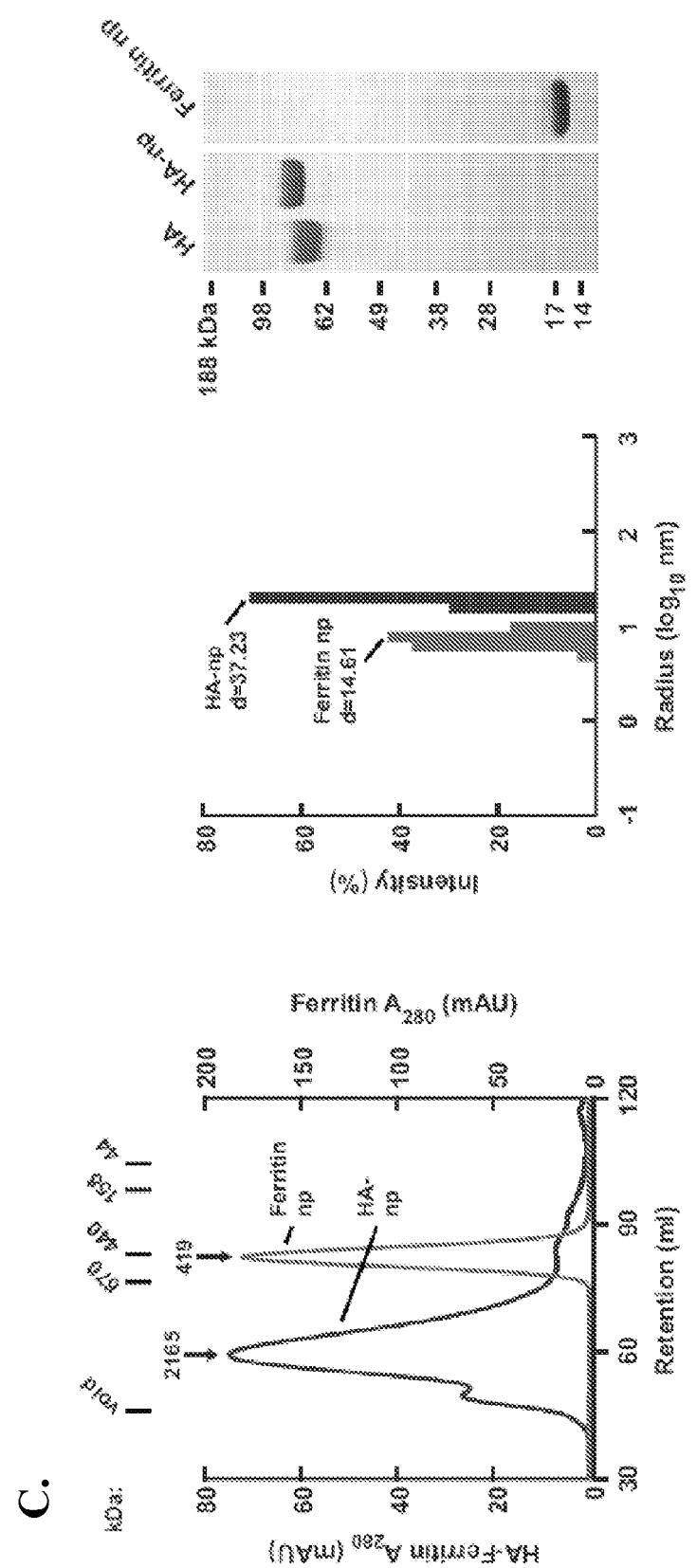


Fig. 1

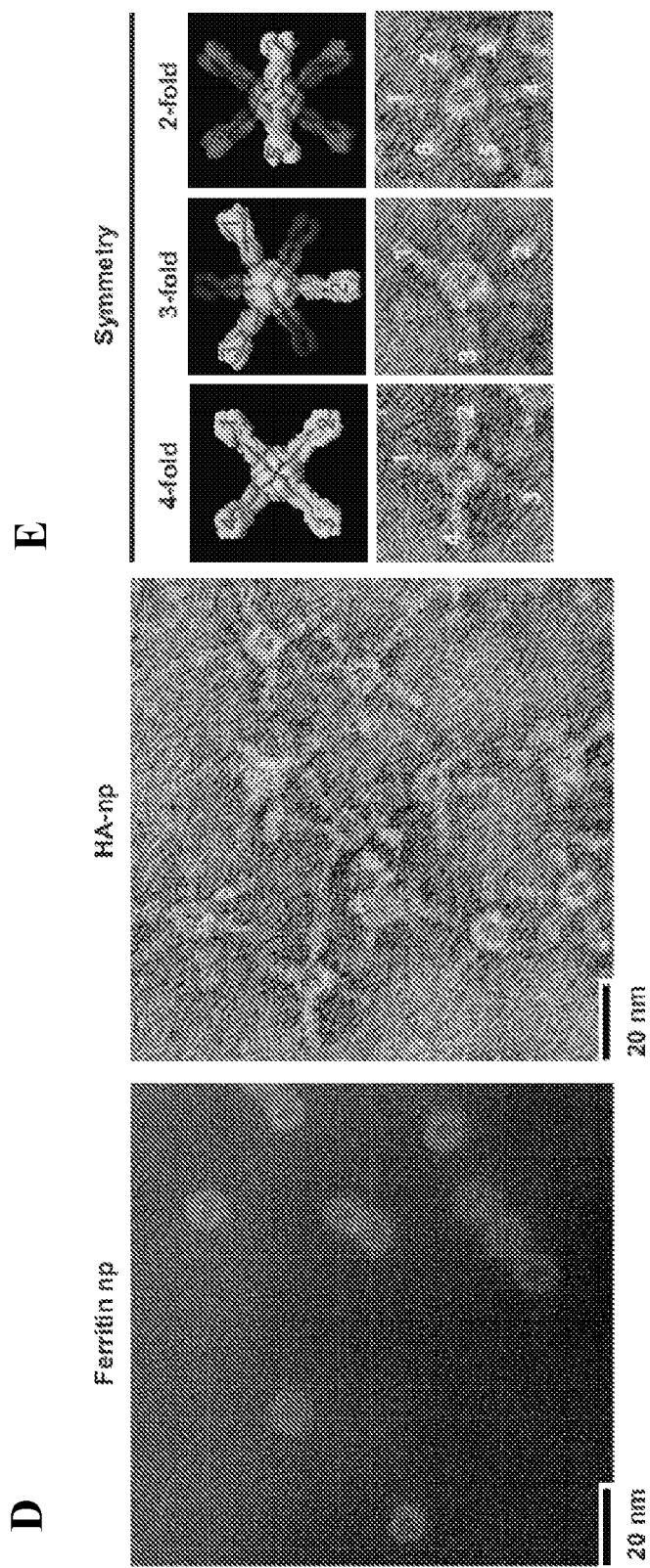
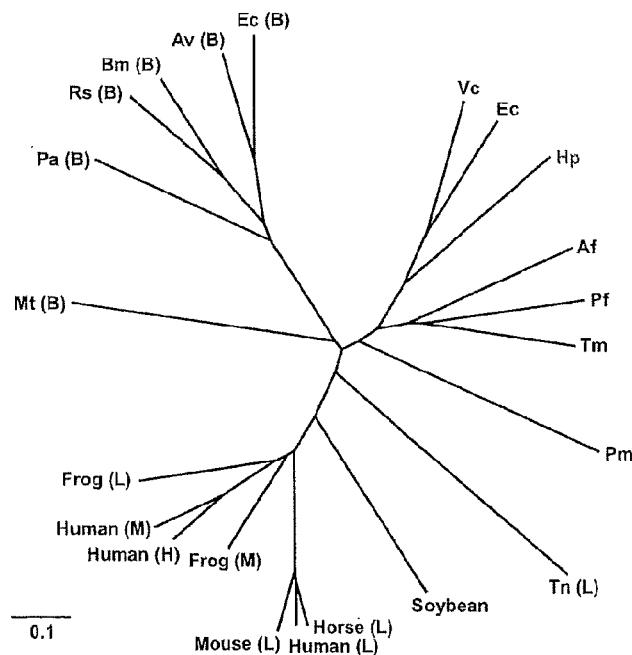
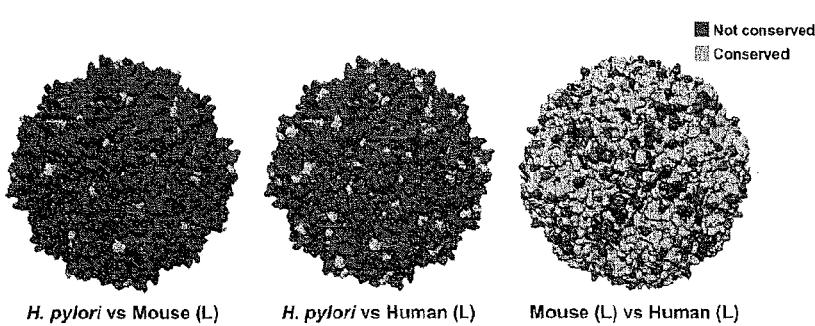
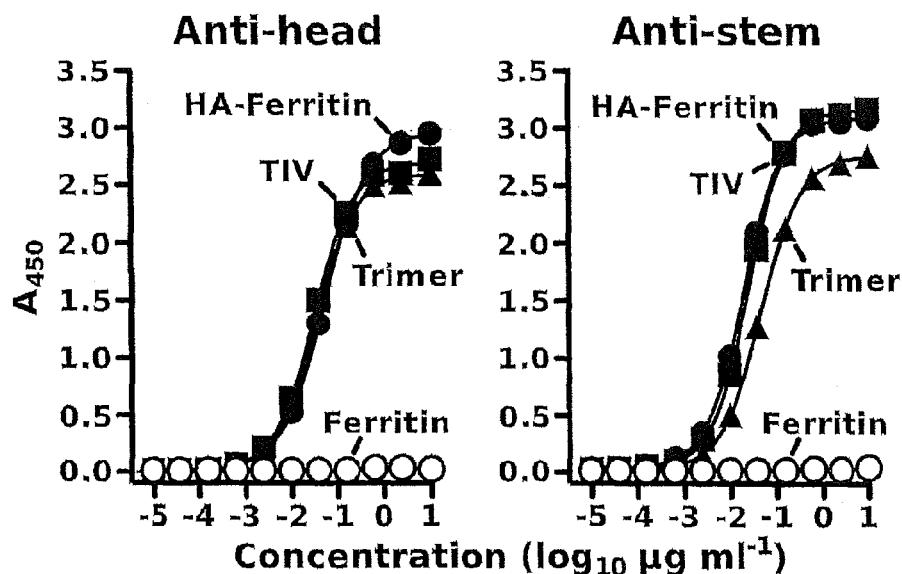


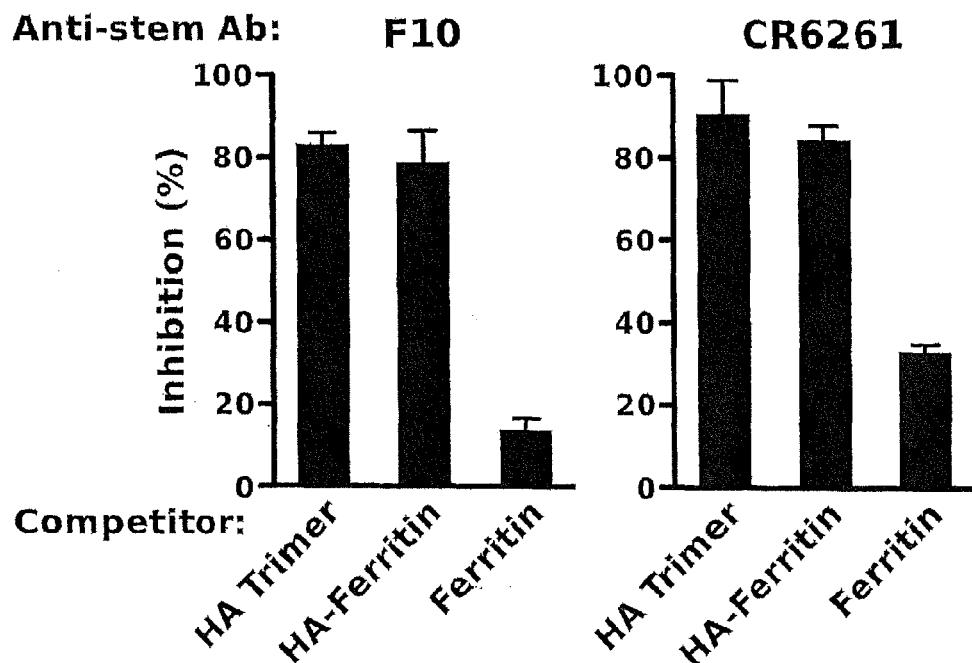
Fig. 1

Supplementary Figure 1 vs2

a**b****Fig. 2**

a

	EC ₅₀ (ng ml ⁻¹)	
	Anti-head	Anti-stem
HA Trimer	34.6	46.6
TIV	31.2	25.4
HA-Ferritin	52.2	19.7

b**Fig. 3**

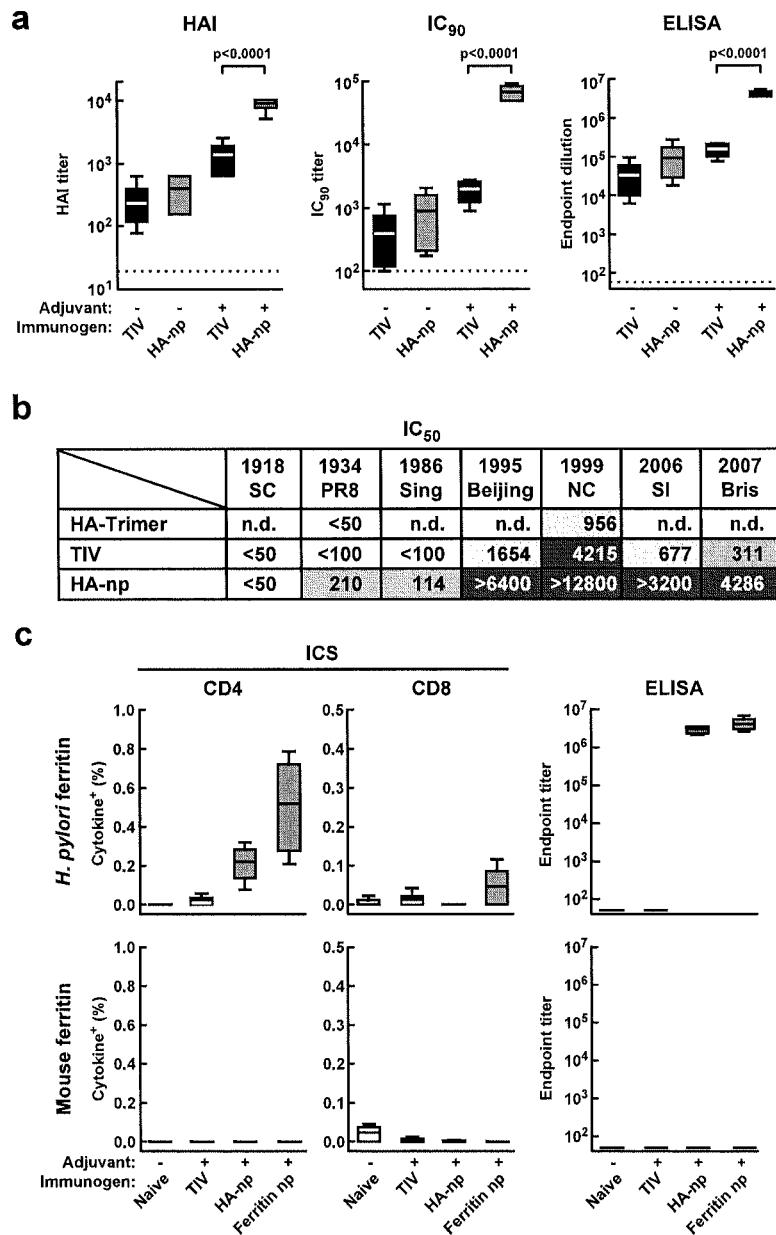


Fig. 4

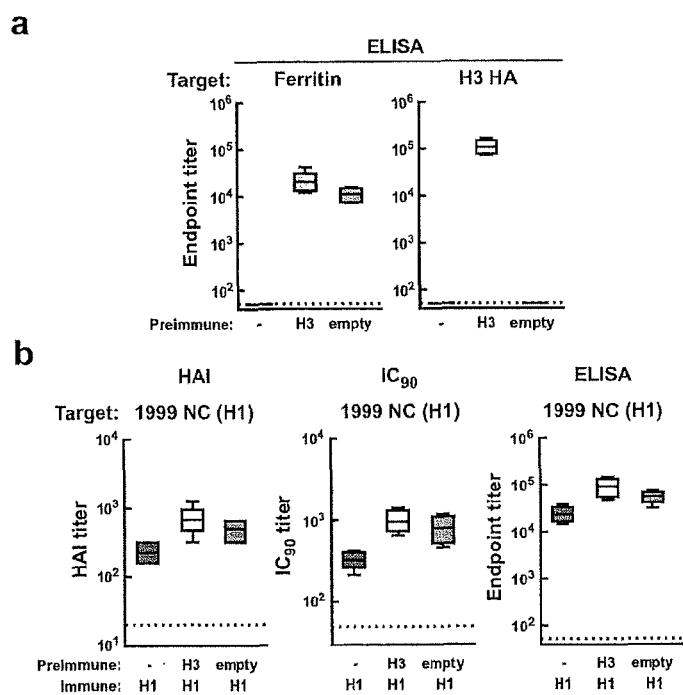
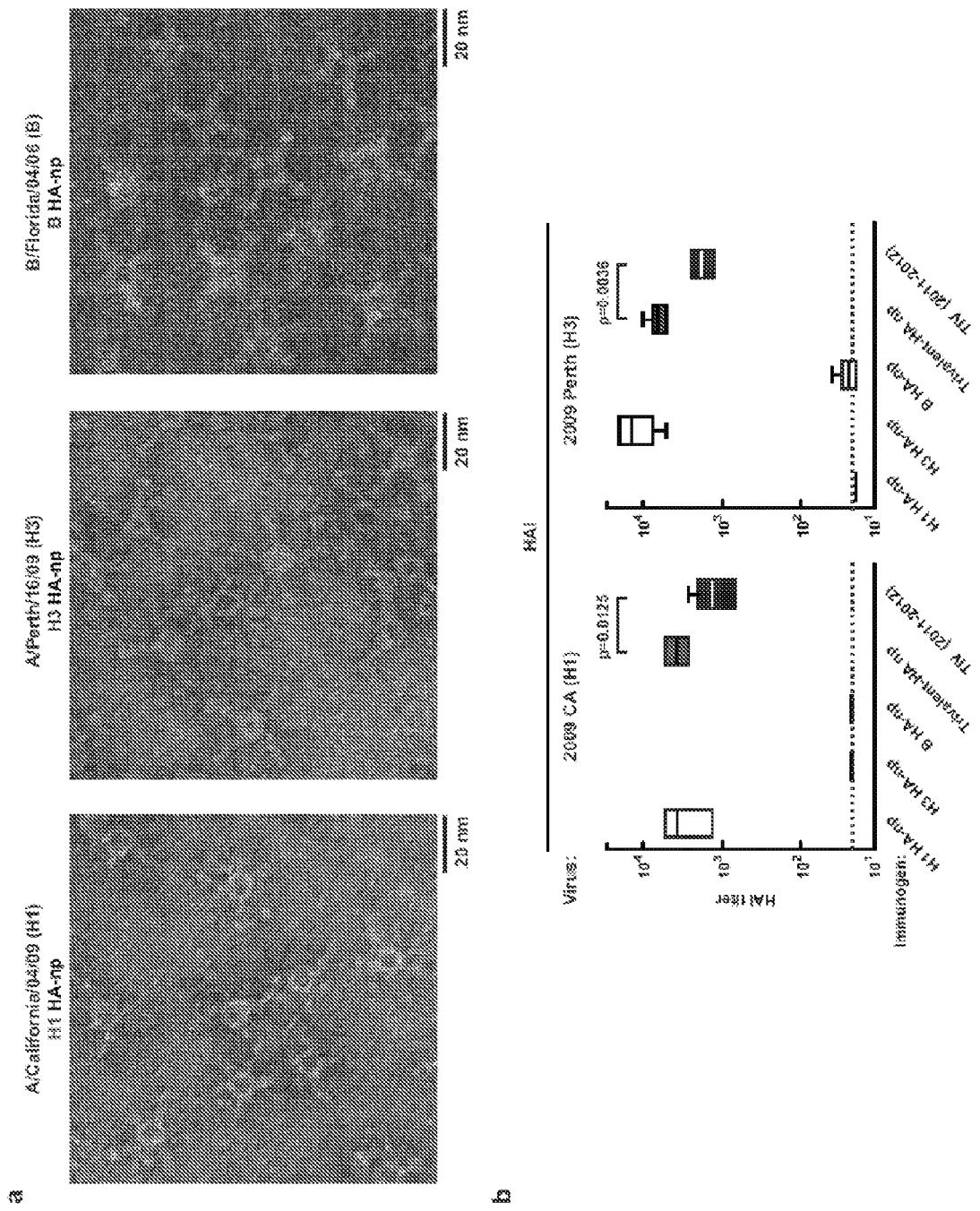


Fig. 5

**Fig. 6**

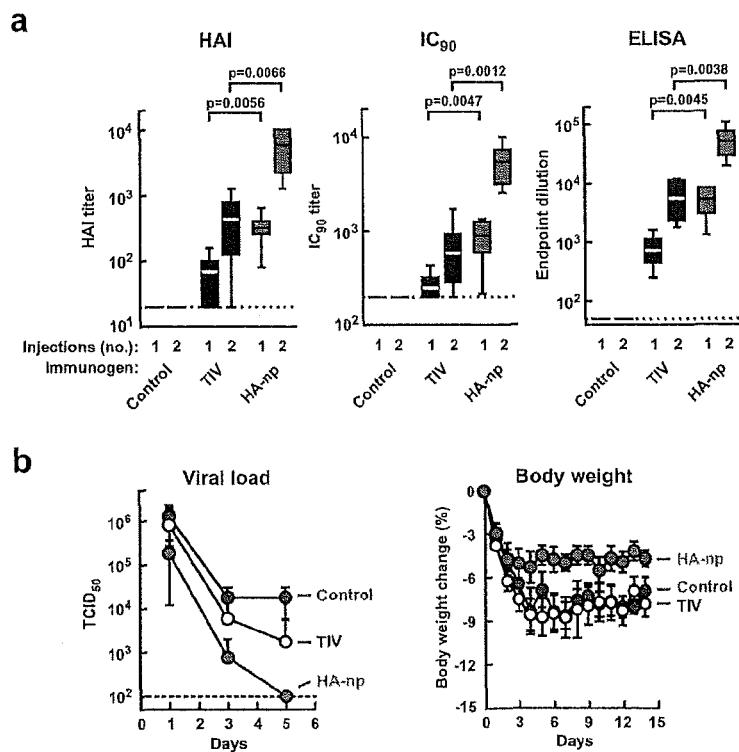


Fig. 7

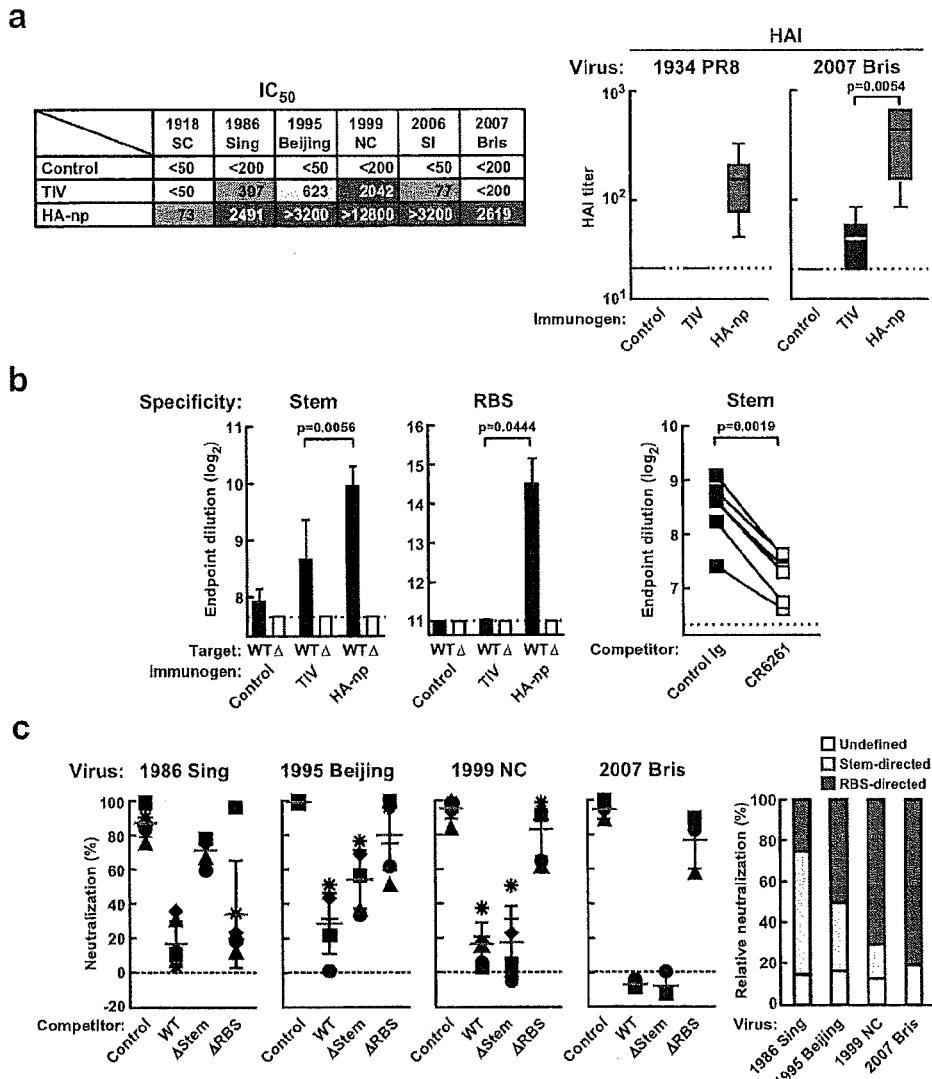


Fig. 8

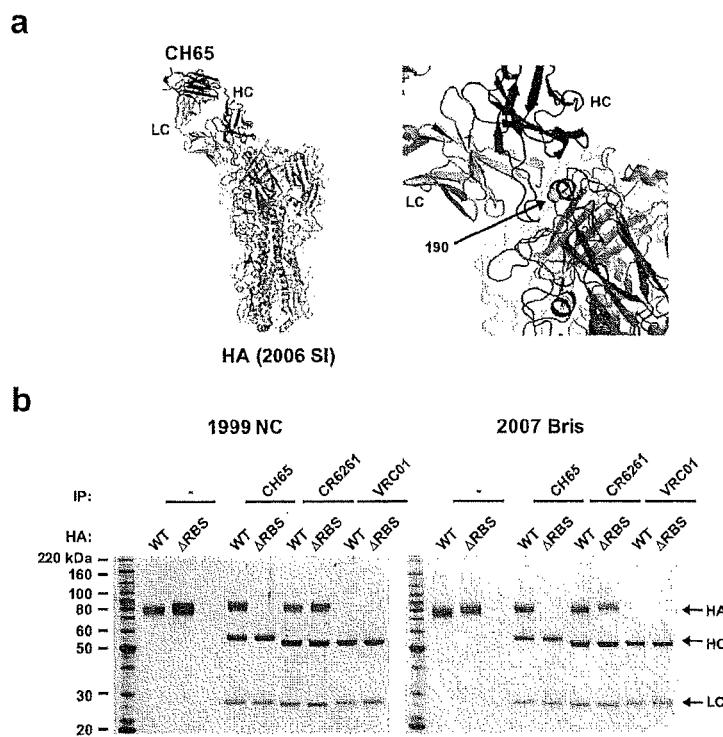


Fig. 9

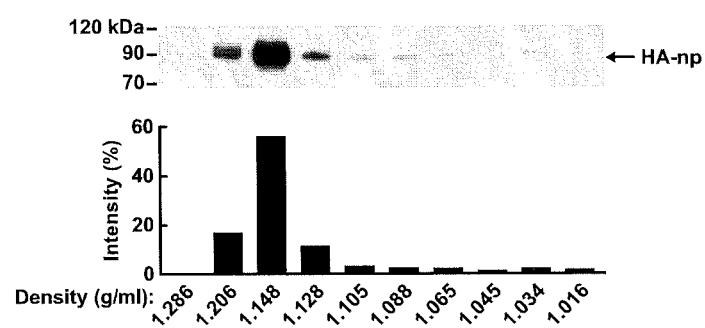


Fig. 10

Immunization of Pan-Group1 HA-Ferritin nanoparticles in Mice and Ferrets

Pan-group1 HA-np vaccine:

H1 A/NC/20/1999
H1 A/CA/04/2009
H2 A/Singapore/1/1957
H5 A/Indonesia/05/2005

Dose:

Mice: 6.8 µg (1.67 µg of each HA-np) + Ribi adjuvant

Ferrets: 10 µg (2.5 µg of each HA-np) + Ribi adjuvant

Injection: Intramuscular, weeks 0 and 3 (mice) or 4 (ferrets)

Fig. 11

ImmunoGenicity of Pan-Group¹ HA-np In Mice

		H1								
Virus	1918 SC	1934 PR8	1947 FM	1954 MAL	1986 SG	1995 BJ	1999 NC	2006 SI	2007 Bris	2009 CA
Animal ID										
5726	2978	<50	<50	<50	<50	<50	8306	690	423	9040
5727	>3200	<50	<50	<50	<50	261	6208	<50	<50	>12800
5728	>3200	<50	76	<50	<50	221	>12800	1360	825	>12800
5729	>3200	<50	<50	<50	<50	989	>12800	<50	121	>12800
5730	2821	<50	<50	<50	<50	4390	10906	<50	<50	11137

		H2							
Virus	1957 SG	2006 MO (Swine)	2007 NED (Avian)						
Animal ID									
5726	>12800	>12800	>12800						
5727	>12800	>12800	>12800						
5728	>12800	>12800	>12800						
5729	>12800	>12800	>12800						
5730	>12800	>12800	>12800						

		H5								
Virus	2004 VN1203	2005 Indo	2006 Nigeria (Avian)	2007 Anhui						
Animal ID	Clade 1	Clade 2.1.3	Clade 2.2	Clade 2.3.4						
5726	488	1128	1438	<50						
5727	100	537	1192	<50						
5728	112	461	1002	<50						
5729	112	335	2539	<50						
5730	105	424	1105	<50						

IC50 titers are shown

Fig. 12

Immunoactivity of Pan-Group1 HA-np In Ferrets

Virus	H1						2009 CA		
	1918 SC	1934 PR8	1947 FM	1954 Mal	1986 SG	1999 NC	2007 Bris		
Animal ID									
456	<50	<50	<50	<50	751	5726	888	>6400	
457	215	<50	474	982	1438	>6400	>6400	>6400	
487	<50	<50	<50	<50	393	3568	402	2417	
488	<50	<50	<50	<50	646	>6400	851	>6400	
492	138	<50	401	<50	1342	>6400	1755	>6400	
493	<50	<50	331	437	815	>6400	3193	>6400	
444	140	<50	<50	<50	421	6400	1223	5547	
445	<50	<50	<50	<50	596	2978	741	3825	
485	<50	<50	<50	<50	501	1962	762	794	
486	<50	<50	50	<50	441	6400	1079	1707	
489	<50	<50	<50	<50	501	4712	455	1172	
490	258	<50	709	885	>6400	>6400	5569	>6400	

Virus	H2			H5				
	1957 SG	2006 MO (Swine)	2007 NED (Avian)	Virus	2004 VN1203	2005 Indo	2006 Nigeria (Avian)	2007 Anhui
Animal ID				Animal ID	Claude 1	Claude 2.1.3	Claude 2.2	Clade 2.3.4
456	2193	247	2351	456	TBD	546	1571	806
457	>6400	1035	>6400	457	TBD	1189	4276	2384
487	2740	284	1637	487	TBD	<50	365	<50
488	>6400	663	>6400	488	TBD	258	900	482
492	>6400	682	>6400	492	TBD	380	1780	731
493	3423	365	3720	493	TBD	335	992	523
444	6400	482	2628	444	TBD	296	817	585
445	762	<50	751	445	TBD	<50	272	<50
485	1425	127	762	485	TBD	<50	<50	<50
486	3329	331	1638	486	TBD	<50	577	112
489	4276	365	2133	489	TBD	50	828	292
490	4276	577	1896	490	TBD	370	1348	449

IC₅₀ titers are shown

Fig. 13

HAI Titters from Pan-Group¹

HA-np Immunized Ferrets

Virus	HAI				
	H1 1986 SG	H1 1999 NC	H1 2007 Bris	H1 2009 CA	H2 1957 SG
Animal ID					
456	16	2560	160	2560	2560
457	40	5120	640	5120	20480
487	40	1280	40	640	2560
488	40	2560	80	1280	10240
492	20	2560	40	5120	10240
493	16	2560	80	5120	2560
444	20	2560	40	1280	2560
445	40	1280	80	1280	640
485	20	1280	20	320	1280
486	20	2560	80	2560	2560
489	16	2560	80	1280	5120
490	40	10240	160	5120	2560

HAI assays were performed using the sera obtained from the immunization studies. The assays were performed using H1N1 viruses or H2 and H5 HA-ferritin nanoparticles.

Fig. 14

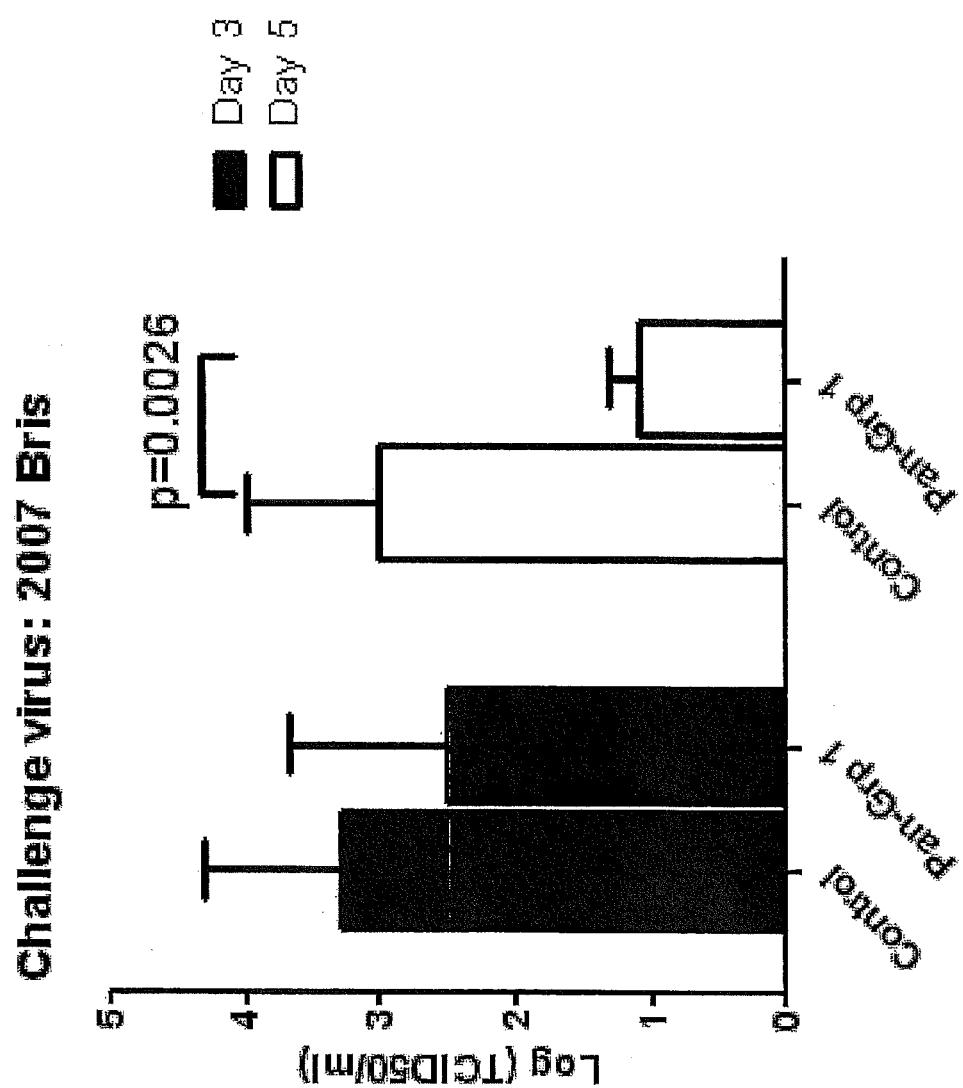


Fig. 15

Challenge virus: 2009 Mex

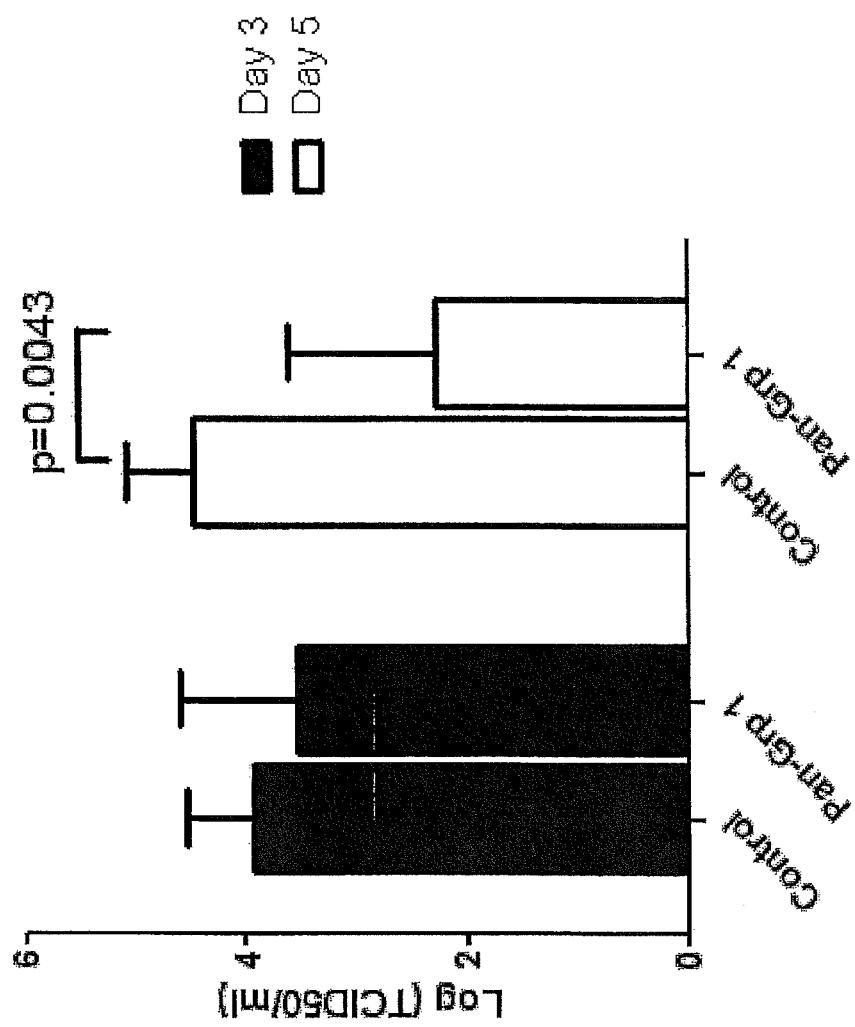


Fig. 16

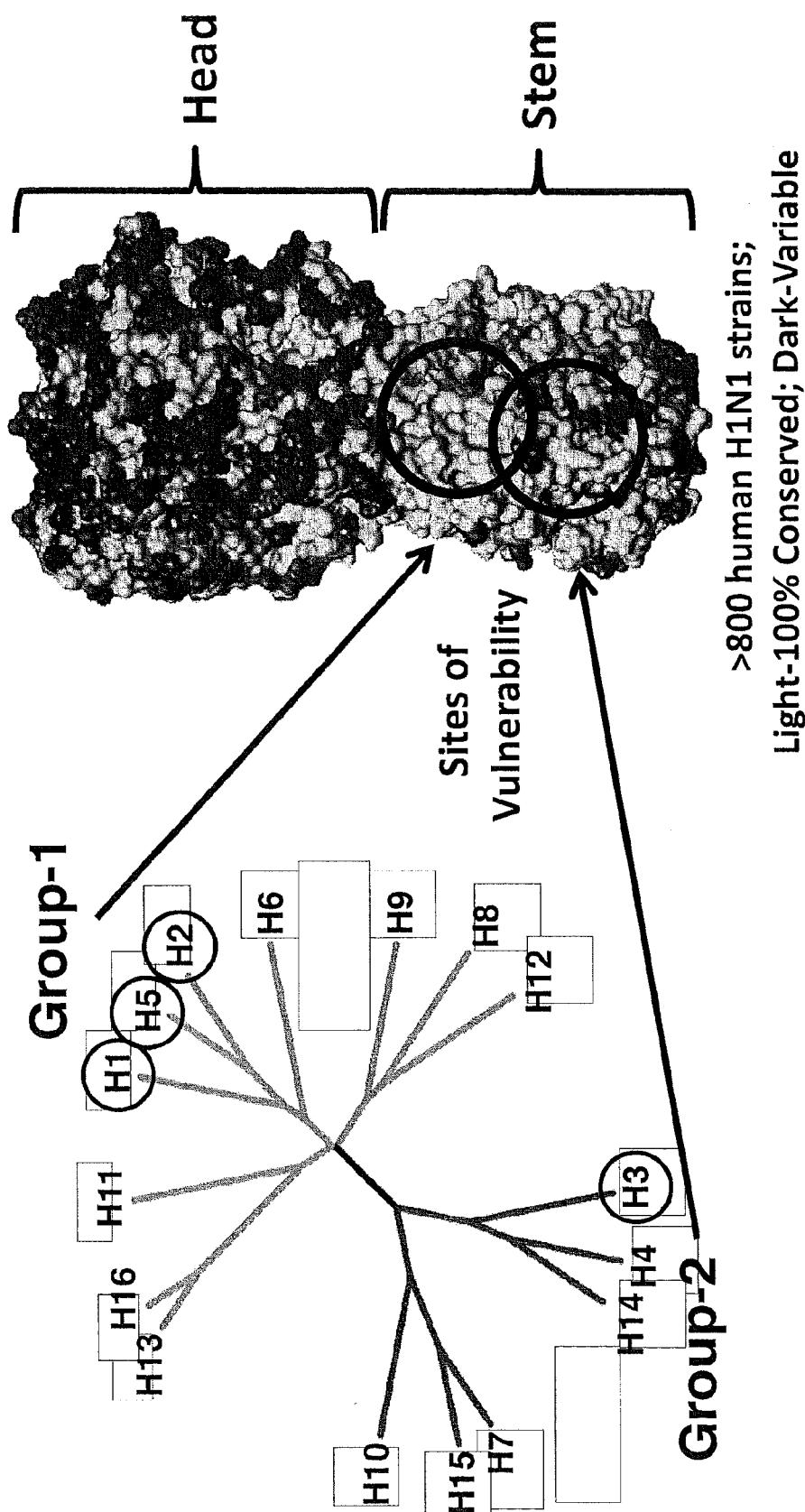


Fig. 17

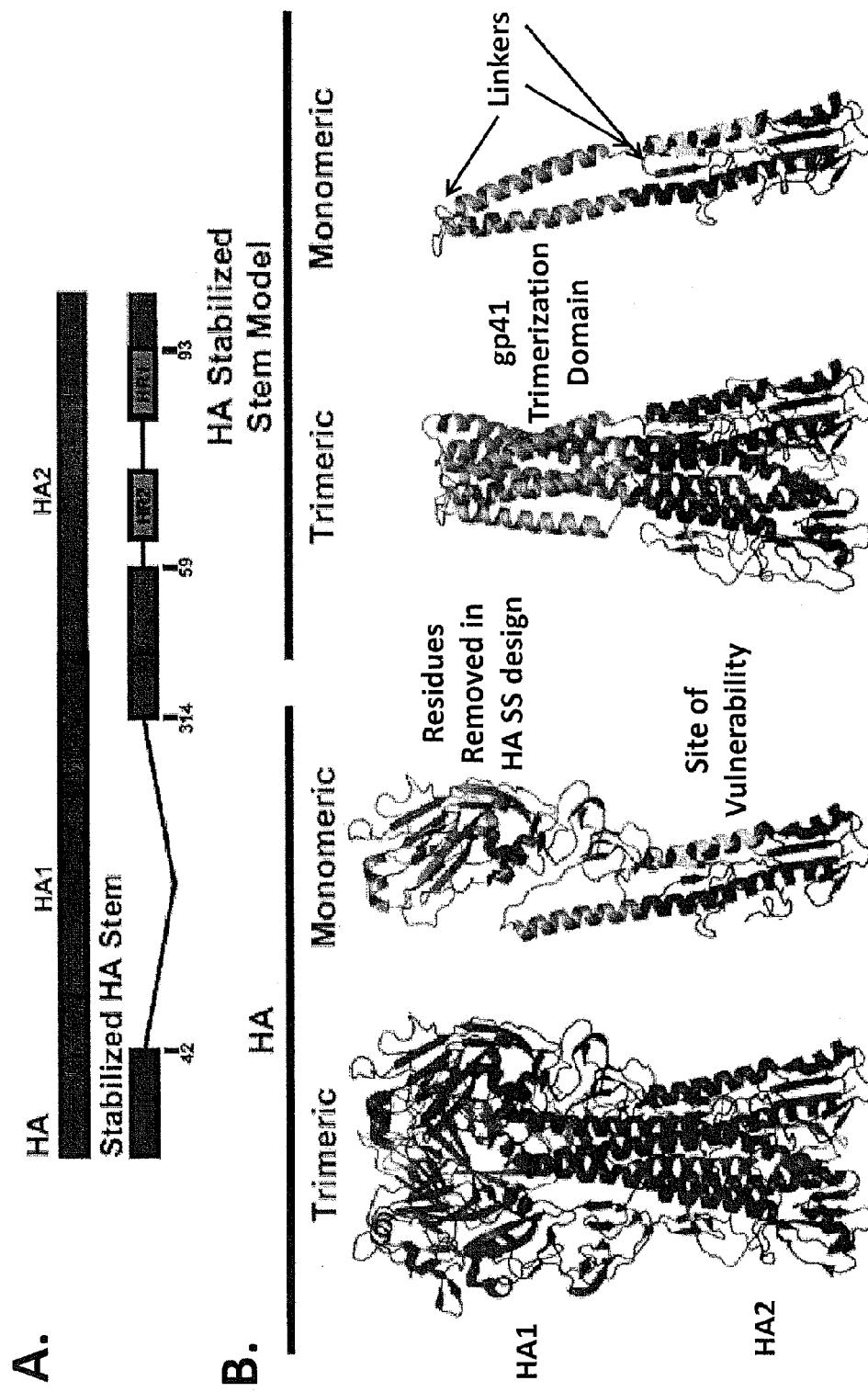


Fig. 18

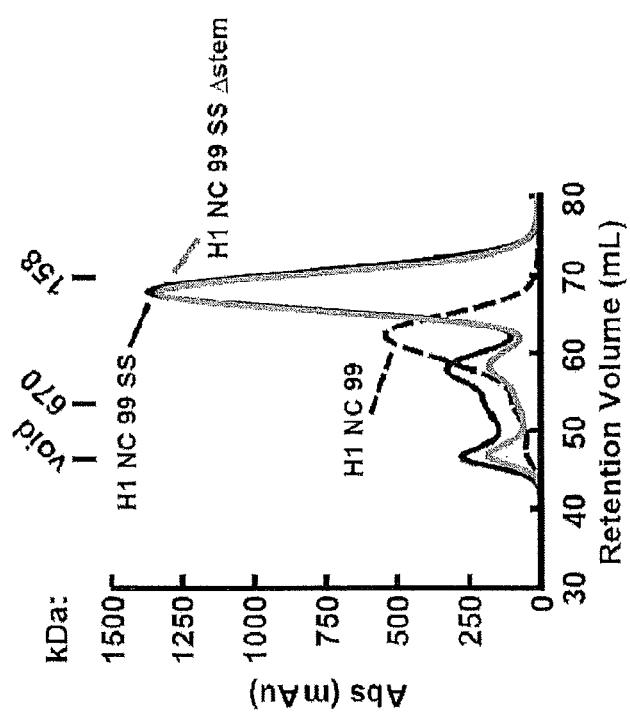


Fig. 19

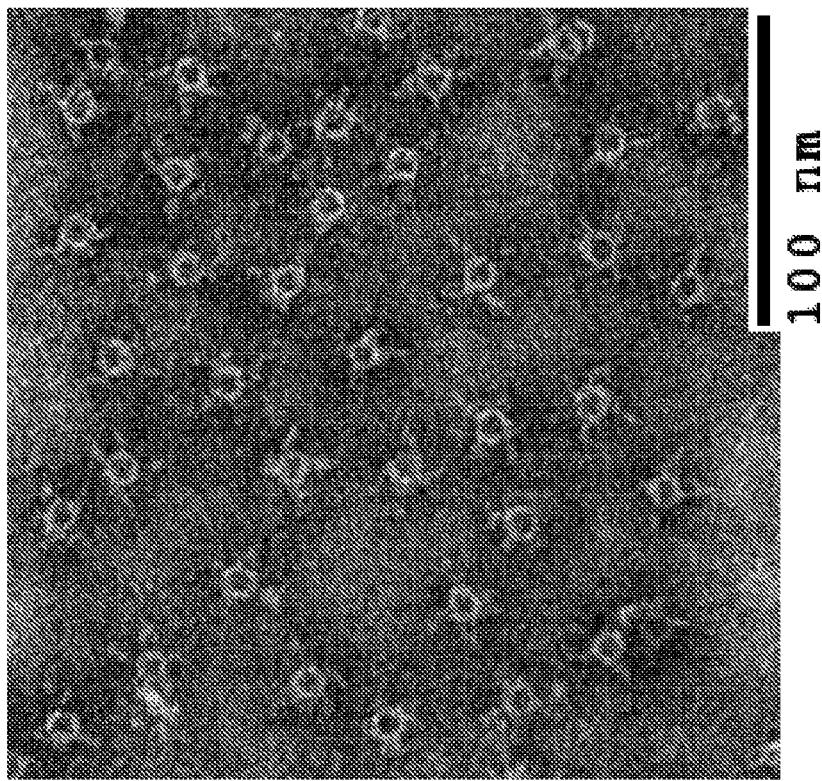
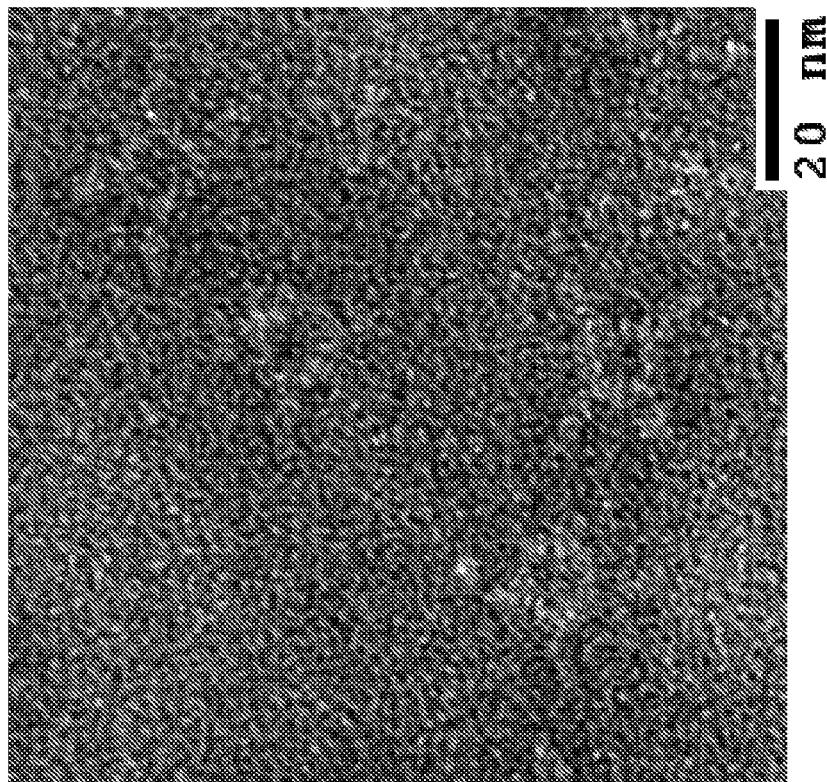


Fig. 20

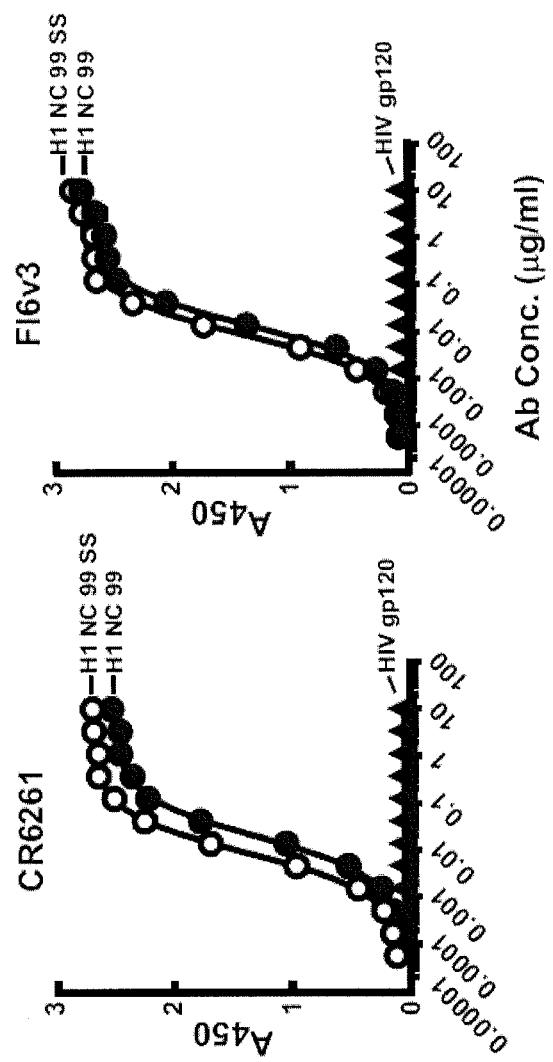


Fig. 21

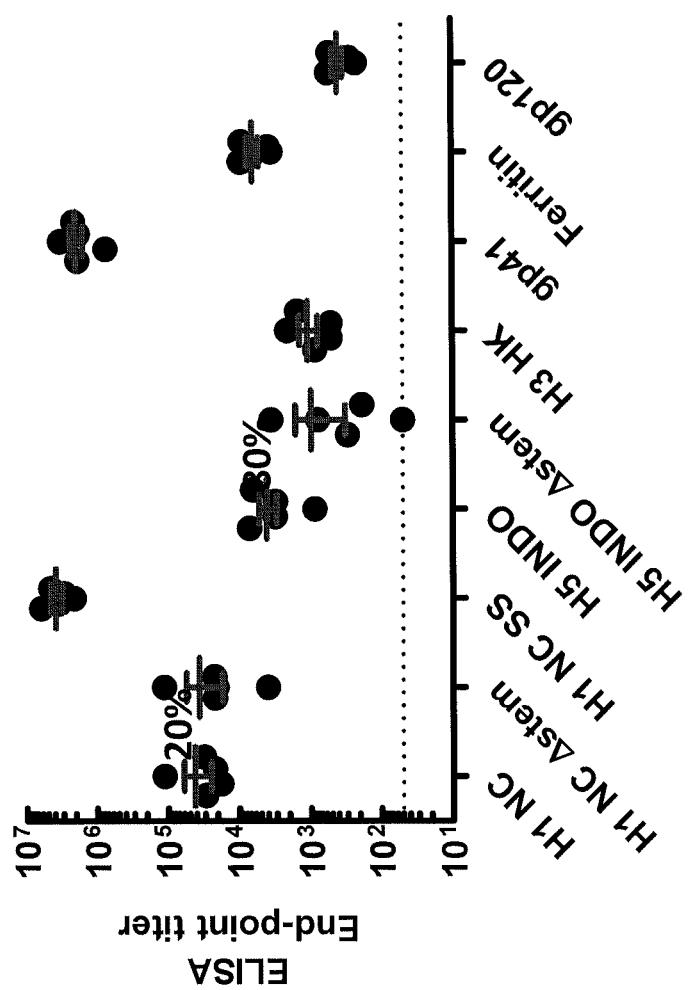


Fig. 22

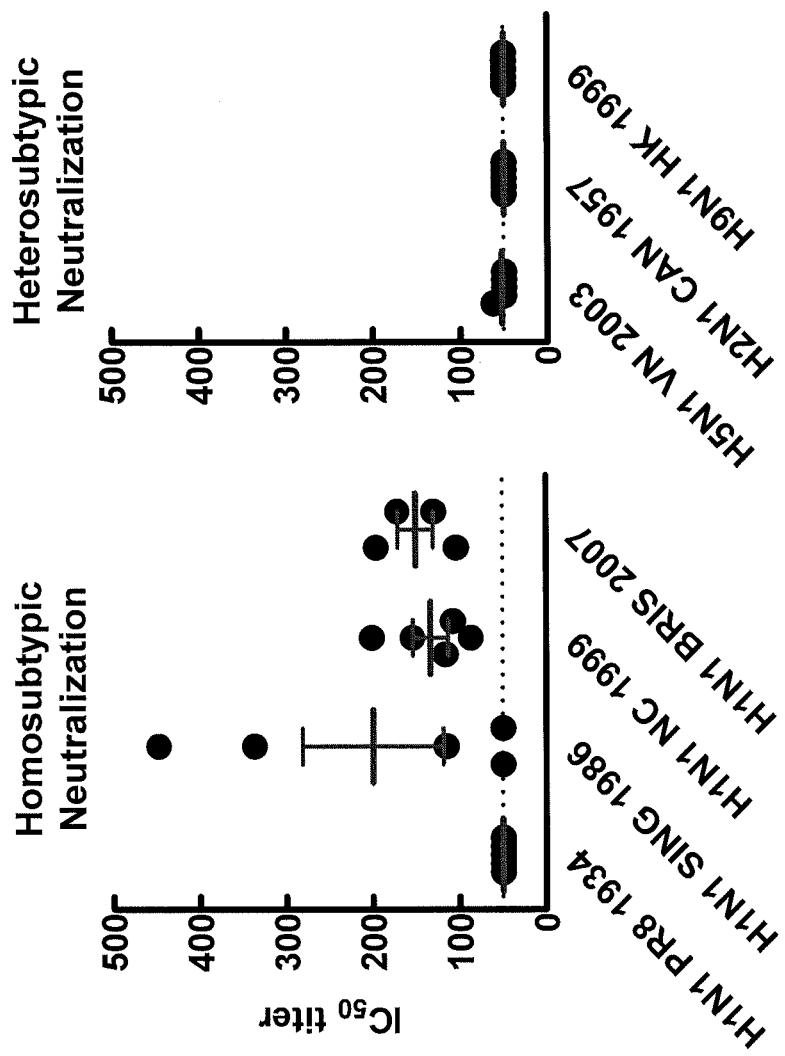


Fig. 23

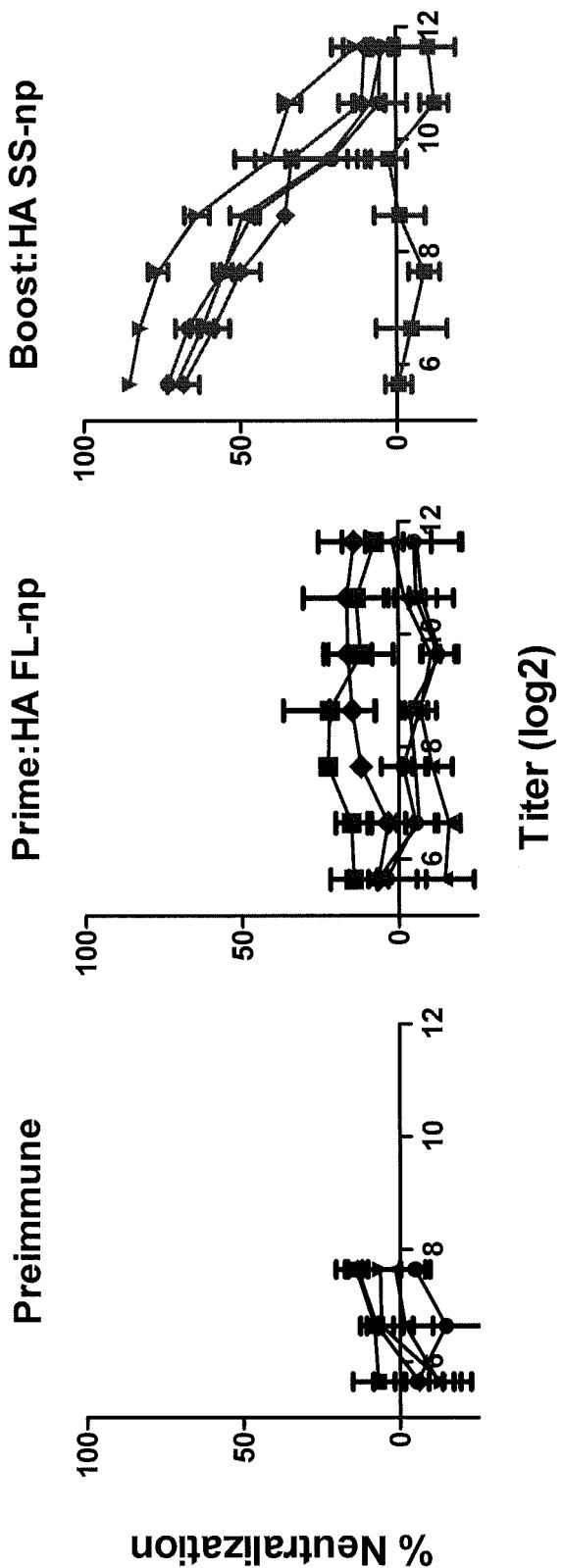
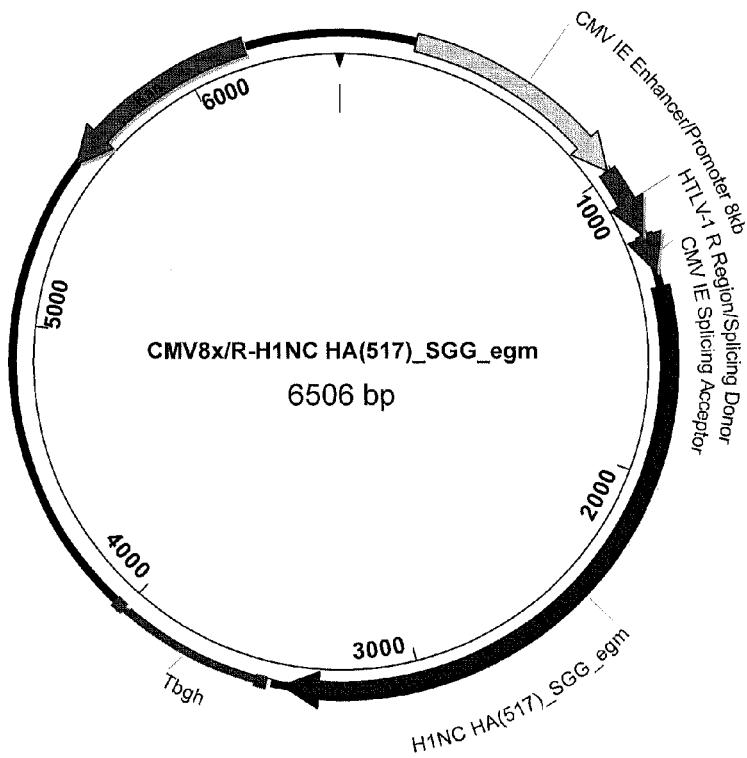


Fig. 24



H1NC HA(517)_SGG_egm (H1 1999NC HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:130)

TCGGCGCTTCCGGTATGACGGTGAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGCTCAGGGGGTGTGGGGGGTGTGGGGCTGGCT
TAACTATGCCCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCCGTGTGAAATACGCCACAGATGCG
TAAGGAGAAAAATACCGCATCAGATTGGCTATTGGCATTGCACTACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACCTCCATAGCCCATAATGGAGTTCCGCTTACATAACCTACGGGAATTTCACAAACCTGGC
TGACCGCCCAACGACCCCCGCCATTGACGTCATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTTGGGAATTCCAAGTGTATCATAT
GCCAACTACGCCCTATTGACGTCAATGACGGGAACCTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGGTTTG

Fig. 25-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCCATGACGTCA
ATGGGAGITTTGTTTGACTCACAAAATCAACGGAAATTCCAAAATGTCGAACAACCTCGCCCCATTGA
CGCAAATGGCGGTAGCGTGTACGGTGGAGGCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCATAGAAGACACGGGACCGATCCACGCCGTTGAGTCGCGTTCTGCC
CGCATCTCTCTTACCGCCCGCCCTACCTGAGGCCCATCCACGCCGTTGAGTCGCGTTCTGCC
GCCTCCCGCCTGCGTGCCTCTGAACCTGCGTCCCGCTTAGCTAAAGCTCAGGTCGAGACCG
GCCTTGCGCCGCTCCCTGGAGGCTACCTAGACTCAGCCGCTCTCCACGCTTGCGTGCCTGACCCCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCTGAGCACTACTCGTGTGCTGCCGCGCGACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCTTCATGGCTTTCTGAGTCACCGTCGACAG
TGTGATCAGATATCGCGCCGCTAGAGATATGCCACCATGAAGGCCAACTGCTGGTGTGCTGTA
CCTTACGCCACCTACCCGACAACTCTGATCGGCTACACGCCAACAAATAGCACCACCGTGGAT
ACAGTGTGGAGAACGTCACCGTGAACCTCTGTGAACCTGCTGGAGGACAGGCCAACATGGCAAGCT
GTGCTGTGAAGGCATTGCCCTCTGAGCTGGCAATTGTTCTGTTGCGGATGGATTCTGGCAACC
CCGAGTGTGAGCTGCTGATTCTAAGAGAGACTGGAGCTACATGTTGAGGCCAACTCTGAGAATGGC
ACCTGCTACCCCTGGTACTTCGCGATTACGAGGAGCTGGCAGCAGCTGCTAGCGTGTCCAGCTTCA
GAGATTCGAGATCTCCCGAAAGGAGTCTGAGCTGGCAGCTGCTAATCACACAGTGTACAGGCGTGTGCTGTA
GCCACACGCCAACAGCTTCTACCGGAACCTGCTGTGGCTGACAGGCAAAATGGCTGACCCCAAC
CTGAGCAAGAGCTAGTGAACAAACAAGGAAAGGAGATGCTGGTGTGGGAGTGCACCCACCCCTCCAA
CATCGGAAATCAGCGGCCCTGTACCCACAGAGAACGCCATGAGCTGGTGTGCTGAGCCACTACAGCA
GAAGATTCAACCCCGAGATCGCAAGAGACCCAAAGTGTAGAGACCCAGGAGGGCCGATCAATTACTACTGG
ACCCCTGCTGGAGCCTGGCGATACCATCATCTCGAGGCCAACCGCAATCTGATGCCCTTGGTATGCCCT
TGCCCTGAGCAGAGGCTTGGCAGCGGCATCATCACAAAGCAACGCCCATGGATGAGTGTGATGCCAACT
GCCAGACACCTCAGGGCGCCATCAATAGCAGCCTGCCCTTCCAGAATGTCACCCCTGTGACCATCGCGAG
TGCCCAAGTATGTGAGAAGCGCCAAGCTGAGAAATGGTACCGGCTGAGAAACATCCCTAGAGGGAGAC
CAGAGGACTGTTGGAGGCATCGCCGATTATCGAGGGAGGATGGACAGGCATGGTGGATGGCTGGTACG
GCTACCACACCAGAATGAGCAGGGCTGGATATGCCGCGATCAGAAGTCTACCCAGAACGCCATCAAC
GGCATCACCAACAAGTGAACAGCGTGTAGCAGAAGATGAACACCCAGTTACCGCTGTGGCAAGGAGTT
CAACAAGCTGGAGCGGAGGATGGAGAACCTGAACAAGAAGGTCGACGCCGTTCTGGACATCTGGACCT
ACAATGCCAACTCTGGCTCTCGAGAATGAGAGGACCCCTGGACTTCCACGACAGAACGTGAAGAAC
CTGTATGAGAAGGCTGAAAGGCCAGCTGAAGAACACAACGCCAAGGGATCGGCTTCCGAGTTCTA
CCCAAGTGAACACAGCTGATGGAGAGCTGAGAAGAACGCCACCTACGACTACCCCTAAGTACAGCGAGG
AGAGCAAGCTGAACGGGAGAAGATGATTGGAGGCGACATCATCAAGCTGCTGAACGGAGCAGGTGAAC
AAGGAGATGCAAGAGCAGAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACAGCTGGACGGCG
CGGCCCTGTCCTGTCACGCGCCGAGGAGTACGAGCACGCCAAGAAGCTGATCATCTTCCCTGAACG
AGAACACGTGCCGCTCAGCTGACCGCATCAGCGCCCCGCCAGCACAGTTCGAGGGCTGACCCAGATC
TTCCAGAAGGCTACGAGCACGAGCACATCAGCGAGAGCATCAACAAACATGTCGACGCCATCAA
GAGCAAGGACCACGCCACCTCAACTTCTGCACTGGTACGGTACGGCCGAGCAGCACGAGGAGGGTGT
TCAAGGACATCCTGGACAAGATCGAGCTGATGCCAACGAGAACACGCCCTGTACCTGGCCACCAGTAC
GTGAAGGCATGCCAACAGCAGGAAGAGCGGATCTAGCATCATCATCATTAGTCTGGAAAGGGCGAA
TTGATCCAGATCTGCTGTGCTTCTAGTTGCCAGCATCTGTTGCTTGCCTCCCTCCCCGTGCTTCTTGA
CCCTGGAAGGTGCCACTCCACTGTGCTTCTAATAAAATGAGGAATTGACATGCAATTGCTGAGTAG
TGTCAATTCTATTCTGGGGGTGGGGCAGGACAGCAAGGGGGAGGATGGGAAGAACATAGCAGGCA
TGGTGGGGATGCCGTTGGGCTCTATGGTACCCAGGTGCTGAAGAATTGACCCGGTTCTCTGGGCCAGAA
AGAAGCAGGCACATCCCTCTGACACACCCCTGTCACGCCACCCCTGGTCTTAGTCCAGCCCCACTC
ATAGGACACTCATGCTCAGGAGGGCTCCCTCAATCCCACCGCTAAAGTACTTGGAGCGGTCTCTCC
CTCCCTCATCAGCCCACAAACCAACCTGCTCAAGAGTGGGAAGAAAATTAAAGCAAGATAGGCTATT
AAGTGCAGAGGGAGGAAAATGCCAACATGTGAGGAAGTAATGAGAGAACATAGAATTAAAGGCC
ATGATTAAAGGCCATCATGCCCTTAATCTCGCTTCTCGCTCACTGACTCGCTCGCTCGTCTCG
CTCGCCGAGCGGTATCGCTCACTAACAGCGGTAAATACGGTTACCCACAGAACATCAGGGATAACGCC
AAAGAACATGTGAGCAAAGGCCAGAAAAGGCCAGGAACCGTAAAAGGCCGCTGCTGGCTTCTCC
ATAGGCTCCGCCCCCTGACGAGCATCACAAATGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAAGATACCAGGCTTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCCCTAC
CGGATACCTGTCCGCCCTTCTCCCTCGGGAAAGCGTGGCGTTCTCATGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTGTTGCTCCAAGCTGGCTGTGACGAACCCCCGGTTCAGGCCGACCGCTGCC

Fig. 25-2

TTATCCGGTAAC TATCGTCTTGAGTCCAACCCGGTAAGACACGA CTTATCGCCACTGGCAGCAGCCACTGG
 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTCTCACAGAGTCTTGAAGTGGTGGCTAACTACGGCT
 ACACAGAAGAACAGTATTGGTATCTGCCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGC
 TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTGCAAGCAGCAGATTACGCGCAG
 AAAAAAGGATCTCAAGAAGATCCTTGTATCTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACAC
 GTTAAGGGATTGCTCATGAGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAAATGAAGT
 TTAAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC
 TATCTCAGCGATCTGTCTATTCGTTCATCCATAGTGCCTGACTCGGGGGGGGGCGCTGAGGTCTGC
 CTCGTGAAGAAGGTGTTGACTCATACCAGGCCTGAATGCCCATCATCCAGCCAGAAAGTGAGGGAG
 CCACGGTTGATGAGAGCTTGTGAGGTGGACCAGTTGGTGAACCTTGCTTTGCCACGGAAACG
 GTCTGCGTTGCGGAAGATGCGTCACTGATCCTCAACTCAGCAAAACTTCGATTATCAACAAAGCC
 GCCGTCGGTCAAGTCAGCGTAATGCTGCCAGTGTACAACCAATTAAACCAATTCTGATTAGAAAAACT
 CATCGAGCATCAAATGAAACTGCAATTATTCATATCAGGATTATCAATACCATATTTTGAAAAGCCGT
 TTCTGTAATGAAGGAGAAAACCTACCGAGGAGCTTCCATAGGATGGCAAGATCTGGTATCGGTCTGCCAT
 TCCGACTCGTCCAACATCAATACACCTATTAAATTTCGTCGTCACAAAGGTTATCAAGTGAGAAAT
 CACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAGCTTATGCATTCCAGACTTGTCAACA
 GGCCAGCCATTACGCTCGTCATCAAATCACTGCATCAACCAACCGTTATTCAATTGTGATTGCGCCTG
 AGCGAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGA
 ACAC TGCCAGCGCATCAAATATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCTGTTTC
 CGGGGATCGCAGTGGTGAACCATGCGATCATCAGGAGTACGGATAAAATGCTTGAATGGTCGGAAGAGG
 CATAAAATTCCGTCAGCCAGTTAGTCTGACCATCTCATCTGTAACATCATTGCCAACGCTACCTTGCCAT
 GTTTCAGAAACAACTCTGGCGCAICGGCTCCCATACAATCGATAGATTGTCGACCTGATTGCCGACA
 TTATCGCGAGCCCATTATACCCATATAATCAGCATCCATGTTGAAATTAAATCGCGGCCCTCGAGCAAGA
 CGTTTCCCGTTGAATATGGCTCATAAACACCCCTGTGATTACTGTTTATGTAAGCAGACAGTTTATTGTC
 ATGATGATATATTTTATCTGTCATGAACTCAGAGATTGAGACACAACGTGGCTTCCCCCCCC
 CCCATTATTGAAGCATTATCAGGGTTATTGTCATGAGGGATACATATTGAATGTTAGAAAAAA
 TAAACAAATAGGGTTCCCGCAGATTCCCGAAAAGTGCCACCTGACGTCAAGAAACCAATTATCA
 TGACATTAACCTATAAAATAGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:131)

ATGAAGGCCAAACTGCTGGTCTGTGTACCTTACCGCCACCTACGCCACACAATCTGTATCGGCTA
 CCACGCCAACAAATAGCACCACCGTGGATACAGTGTGCTGGAGAAGAACGTGACCGTGAACCGTAACTCTGTA
 ACCTGCTGGAGGACAGCCACAATGGCAAGCTGTGTGCTGAAAGGCATTGCCCTCTGCAGCTGGCAAT
 TGTCTGTGGCCGGATGGATTCTGGCAACCCCGAGTGTGAGCTGCTGATTCTAAGGAGAGCTGGAGCTA
 CATCGTGGAGACCCCCAATCTGAGAATGGCACCTGCTACCCCTGGCTACTGCCGATTACGAGGAGCTGC
 GCGAGCAGCTGTCTAGCGTGTCCAGCTCGAGAGATTCGAGATCTCCCAAGGAGTCCAGCTGGCTTAAT
 CACACAGTGACAGCGTGTGCTGCCAGCTGTAGCCACAACGGCAAAGCAGCTTCTACCGAACCTGCTGTG
 GCTGACAGGCAAGAATGGCTGTACCCCAACCTGAGCAAGAGCTACGTGAAACAACAGGAAAGTGC
 TGGTCTGTGGGAGTGCACCACCCCTCCAACATCGGAAATCAGGGCCCTGTACACACAGAGAACGCC
 TATGTGAGCGTGGTGTCCAGCCACTACAGCAGAAAGATTACCCCCCGAGATGCCAAGAGAACCCAAAGTGA
 AGACCAAGGAGGGCCGGATCAATTACTACTGGACCTGCTGGAGCTGGCGATACCATCATTCGAGGCCA
 ACGGAATCTGATGCCCTGGTATGCCCTGGTACGGAGGCTTGGCAGCGGCATCATCACAAGC
 AACGCCCATGGATGAGTGTATGCCAAGTGCAGACACCTCAGGGGCCATCAATAGCAGCTGCCCTT
 CCAGAAATGTGACCCCTGTGACCATCGGGAGTCCCCAAGTATGTGAGAAGGCCAAGCTGAGAATGGTGA
 CGGGCCTGAGAAACATCCCTCAGAGGGAGACCGAGGAGACTGTTGGAGGCCATGCCGGATTCTGAGGGA
 GGATGGACAGGCATGGTGGATGGCTGGTACCGCTACCACCCAGAATGAGCAGGGCTCTGGATATGCCG
 CGATCAGAAAGCTACCCAGAACGCCATCACCGCATACCAACAAGGTGAACACGGCTGATCGAGAAGATGA
 ACACCCAGTTACCGCTGTGGCAAGGGATTCAACAAGCTGGAGCGGAGGATGGAGAACCTGAACAAGAAG
 GTGGACAGCGCTTCTGGACATCTGGACCTACAATGCCAACTCTGGTCTCTGAGAATGAGAGGAC
 CCTGGACTTCCACGACAGCAACGCTGAAGAACCTGTATGAGAAGGTGAAGAGGCCAGCTGAAGAACACGCCA
 AGGAGATCGGAACGGCTGTTGAGTTCTACACAAAGTGAACACAGGTGATGGAGAGCGTGAAGAAC
 GGCACCTACGACTACCCCTAAGTACAGCGAGGAGAGCAAGCTGAACCGGGAGAAGATGAGGCCA
 CATCATCAAGCTGCTGAACGAGCAGGTGAACAAGAGATGAGCAGAGCAGCAACCTGTACATGAGCATGAGCA
 GCTGGTCTACACCCACAGCCTGGACGGCGCCCTGTTCTGACCACGCCAGGGAGTACGAG

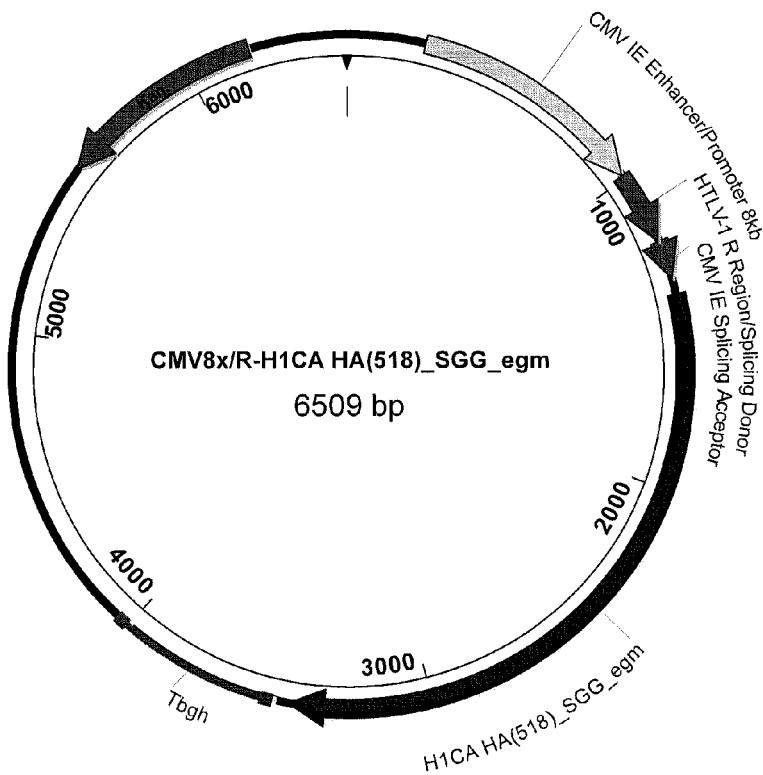
Fig. 25-3

CACGCCAAGAACGCTGATCATCTTCCTGAACGAGAACACGTGCCGTGCAGCTGACCAGCATCAGCGCCCC
CGAGCACAAGTTCGAGGGCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCACATCAGCGAGA
GCATCAACAAACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCTGCAGTGGTAC
GTGGCCGAGCAGCACGAGGAGGTGCTGTTCAAGGACATCCTGGACAAGATCGAGCTGATCGGCAACGA
GAACCACGGCCTGTACCTGGCCGACCAGTACGTGAAGGGCATCGCCAAGAGCAGGAAGAGCGGATCCTAG

Translation (SEQ ID NO:41)

MKA
KLLVLLCTFTATYADTICIGYHANNSTD
TVDTVLEKNVTVTHSVNLLED
SHNGKLCLLKGIAPLQLGN
CSVAGWILGNPECELLISKE
SWSYIVETPNPENG
TCYPGYFADYEELREQ
LSSVSSFERFEIFPKES
SWPN
HTVTGVSA
CSHNGKSSFYRNLLWLTGKNGLYP
NLSKS
YVNNKEKEV
LVLWG
VHHPPNIGNQ
RALYHT
ENA
YVS
VVSSHY
SRRTPEIA
KRP
KVRD
QEGR
INY
WT
LLEPG
GDTI
IFE
ANG
NLIA
PWY
AFAL
SRGF
GSGI
ITS
NAPM
DEC
DAK
CQTP
PQ
GAIN
SSL
PFQ
QN
VHP
VTIG
ECP
KY
VRS
AKL
RMV
TGL
RNIP
QRE
TRGL
FGAI
AGF
IEG
GWTGM
VDGW
YHH
HQNE
QGS
GYA
ADQ
KST
QNA
ING
ITN
KV
NSV
IEK
MNT
QFT
AVG
KEF
NKL
ERR
MEN
LNKK
VDD
GFL
DIW
TYNA
ELL
VL
LEN
ERT
LDF
HDS
NV
KNL
YE
KV
SQL
KNN
AKE
I
NGC
FE
FY
HK
CN
NE
CM
EV
KN
GTY
DYP
PKY
SE
ES
KLN
RE
KID
SGG
DI
IK
LL
NE
QVN
KEM
QSS
NLY
MS
MSS
CY
HSL
DG
AGL
FL
FD
HAA
EYE
HAK
KLI
I
FLN
ENN
NP
VQL
TS
ISA
PEH
KF
EGL
TQ
IF
Q
KAY
E
HQ
H
I
SES
IN
N
I
V
D
HAI
KS
KD
HAT
FN
FL
QW
Y
VA
EQ
HE
EE
EV
LF
KD
ILD
KIEL
IGN
EN
HGL
Y
LA
DQ
YV
KG
IA
KS
RK
SGS

Fig. 25-4



H1CA HA(518)_SGG_egm (H1_2009CA HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:132)

```

TCGGCGCTTCCGGTATGACGGTAAACCTCTGACACATGCGAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAAGCGGGTGTGGCGGGTGTGGGGCTGGCT
TAACTATGCGCATCAGAGCAGATTGACTGAGAGTGACACCATATGCGGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGAATTGGCTATTGGCATTGCATACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACCGCCATTGACATTGATTATTGACTAGTTATAATAGTAATCA
ATTACGGGAACCTCCATAGCCCCATATATGGAGTTCCCGCTTACATAACTTACCGGAATTCCAACCTGGC
TGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCATGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCTATTGACGTCAATGACGGAACTTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGGTTTG

```

Fig. 26-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTTGTTTGTACTCACAAAATCAACGGGAATTCCAAAATGTCGTAACAACCTCGCCCCATTGA
CGAAATGGCGGTAGCGTACGGGGAGGTCTATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACCGTGTGTTGACCTCCATAGAAGACACCGGACCGATCCAGCCCTCCATCGGCT
CGCATCTCTCTTCACGCCCGCCGCCATCTGAGGCCATCCACGCCGGTTGAGTCGCGTTCTGCC
GCCTCCCGCCTGTTGCTGCTGACTCGCTCCGGCTAGGTAAGTTAAAGCTCAGGTGAGACCCGG
GCCTTGCTCCGGCGCTCCCTGGAGCTACCTAGACTCAGCCGGCTCTCCACGCTTGCTGACCCCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGACTGTAGTCTGAGCAGTACTCGTTGCTGCCGCCACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCCCTTCATGGTCTTCTGAGTCACCGTCGACACG
TGTGATCAGATATCGCGGCCGCTCTAGAGATATCGCCACCATGAAGGCCATCTGGTGGTGCTGCTGTACA
CCTTCGCCACCGCAACGCCACCCCTGTGATCGCTACCCAGCCAACACAGCACCGACACCGTGGAC
ACCGTGCTGGAGAAGAACGTGACCGTACCCACAGCGTAAACCTGCTGGAGGACAAGCACAACGGCAAGCT
GTGCAAGCTGCCGGCGTGGCCCCCTGCACCTGGCAAGTGCACACATGCCGGCTGGATTCTGGCAACC
CCGAGTGCAGAGCCCTGAGCACCGCAGCAGCTGGAGCTACATGTTGAGACCCCCAGCAGCGACAACGGC
ACCTGCTACCCCGCGCACTTCATGACTACGAGGAGCTGCCAGCTGAGCAGCGTGAGCAGCTTCGA
GCCGTTGAGATCTCCCCAAGACCAGCAGCTGGCCAACCACAGCAACAAGGGCGTGACCGCCGCT
GCCCGCACGCCGGCGCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCC
AAGCTGAGCAAGAGCTACATCACGACAAGGGCAAGGAGGTGCTGGTGTGGGGCATCCACCCACCCAG
CACCGCGCCGACCAGCAGAGCTGTACCGAGAACGCCACCTACGTTGCTGGCAGCAGCGGTACA
GCAAGAAGTCAAGCCCAGAGTCGCCATCGGCCAACGGTGCAGGACAGGAGGGCGGATGAACTACTAC
TGGACCTCTGGAGGCCCGCAAGATCACCTGAGGCCACCGGCAACCTGGTGGTGGCCGTTACGC
CTTCGCCATGGAGCGGAACGCCGCGAGCCATCATCACGAGCACACCCCGTGCACGACTGCAACACCA
CCTGCCAGACCCCAAGGGGCCATCAACACAGGCTGCCAACGGTGCAGGACATCCACCCATCACCATCGGC
AAGTCCCCAAGTACGTGAAGAGCACAAGCTGCCGCTGGCCACCGGCTGGGAACATCCCCAGCATCCA
GAGCCGGGGCTGTCGGGCCATCGCCGCTTCATCGAGGGCGCTGGACCGGATGGTGGACGGCTGGT
ACGGCTACCCACCAAGAGCTGGAGAAGCGGATCGAGAACCTGAACAAGGGTGGACGGCTTCTGGACATCTGGA
GTCACACCACCTGGAGAAGCGGATCGAGAACCTGAACAAGGGTGGACGGCTTCTGGACATCTGGA
CTTACAACGCGAGCTGCTGGTGTGGAGAACAGGAGCTTACACGACAGCAACGTGAAG
AACCTGTAAGAGAAGGTGCGGAGCCAGCTGAAGAACACGCCAAGGGAGATCGCAACGGCTGCTGAGTT
CTACCAAGTGCACACACCTGCATGGAGAGCGTGAAGAACGGCACCTGACTACCCAAAGTACAGCG
AGGAGGCCAACGCTGAACCGGGAGGAGATCGACTCCGGAGGCACATCATCAGCTGCTGAACGAGCAGGTG
AACAAAGGAGATGCAAGCAGCAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACGCTGGACGG
CGCCGGCCTGTTCTGTCGACCAACGCCGAGGAGTACGAGCAGGCCAACAGCTGATCATCTTCTG
ACGAGAACACGTCGCCGTGAGCTGACCGCATCAGGCCCGGAGCACAAAGTCTGAGGGCTGACCCAG
ATCTCCAGAAGGCCAACGAGCAGCACATCAGCGAGAGCATCAACACATCGTGGACCAACGCCAT
CAAGAGCAAGGACCACGCCACCTCAACTTCTGCACTGAGCTGGTACGTGGCAGCAGCACGAGGAGGTG
TGTTCAGGACATCTGGACAAGATCGAGCTGATCGCAACGAGAACACGCCCTGTACTGGCGACCGAG
TACGTGAAGGGCATGCCAACAGCAGGAAGAGCGGATCCTAGCATCATCATCATATTAGTCTGGAAAGGGC
GAATTGATCCAGATCTGCTGTGCTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCTTCCT
TGACCCCTGGAGGTGCCACTCCACTGTGCTCTTCTTAATAAAATGAGGAATTGCATGCCATTGCTGAGT
AGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
GCATGCTGGGGATGCGGGGGCTCTATGGGTACCCAGGTGCTGAAGAATTGACCCGGTTCTCTGGCCA
GAAAGAACGGCACATCCCTCTCTGTGACACACCCCTGTCCACGCCCTGGTCTTAGTCCAGCCCCA
CTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCATCCACCCACCGCTAAAGTACTTGGAGCGGTCTC
TCCCTCCCTCATGCCAACCAAACCTAGCCTCCAAGAGTGGGAAGAAATTAAAGCAAGATAGGCT
ATTAAGTGCAGAGGGAGAGAAAATGCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAATTAAAG
GCCATGATTTAAGGCCATCATGGCTTAATCTCCGCTCTCGCTACTGACTCGCTGCCCTGGTCTGTT
CGGCTGCCGAGCGGTATCAGCTCACTAACGGGCTAAAGGTTATCCACAGAAATCAGGGATAACGC
AGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAGGCCGCTGGCTGGCTTT
TCCATAGGCTCCGCCCTGACGAGCATCACAAAAATGACGCTCAAGTCAGAGGTGGCGAACCCGACA
GGACTATAAGATACCAGCGTTCCCTGGAGGCTCCCTCGTGCCTCTCTGTTGCCACCCCTGCCGCT
TACCGGATACCTGTCGCCCTTCTCCCTGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATC
TCAGTCGGTGTAGGTGCTGCCAACGCTGGCTGTGCAAGAACCCCCGGTTCAGGCCGACCGCTGC

Fig. 26-2

GCCTTATCCGTAACATCGTCTTGAGTCCAACCCGTAAGACACGACTTATGCCACTGGCAGCAGCCAC
TGGTAACAGGATTAGCAGAGCAGGTATCTAGGGTGCTACAGAGTTGAAGTGGTGGCTAACCTACG
GCTACACTAGAAGAACAGTATTGGTATCTGCCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
AGCTCTTGATCCGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTGTGCAAGCAGCAGATTACCG
CAGAAAAAAAGGATCTCAAGAAGATCCTTGTATCTACGGGCTGACGCTCAGTGAACGAAACT
CACGTTAAGGGATTGGTATGAGATTATCAAAAGGATCTCACCTAGATCCTTTAAATTAAAATGA
AGTTTAAATCAATCTAAAGTATATGAGTAACTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTGCTCATCCATAGTTGCTGACTCGGGGGGGGGCGCTGAGGTC
TGCCTCGTGAAGAAGGGTGTGACTCATACCCAGGCTGAACTGCCCATCATCCAGCAGAAAGTGAGG
GAGCCACGGTGTGAGAGCTTGTGAGTGGGACAGTTGCTGATTTGAACTTTCTTGCACCGA
ACGGTCTGCTGGTGTGGAGATGCGTATCTGATCTGATCCTCAACTCAGCAAAGTGATTTATTCAACAAA
GCCCGCGCCCGTCAAGTCAGCGTAATGCTCTGCGAGTGTACAACCAATTAAACCAATTCTGATTAGAAAA
ACTCATGAGCATCAAATGAAACTGCAATTATTATCATCAGGATTATCATACCATATTTGAAAGAAC
CGTTTCTGTAATGAAGGGAGAAACTACCCAGGAGCTTCCATAGGATGGCAAGATCCTGGTATCGTCTGC
GATTCCGACTCGTCAACATCAATACAACCTATTAAATTCCCGTCAAAATAAGGTTATCAAGTGAGA
AATCACCAGTGAAGTGAAGTCAACTGAGTAACTGGTGGAGAATGGCAAAAGCTTATGCTATTCTTCCAGACTTGGTCA
ACAGGCCAGGCAATTACCGCATCGTCAATCAAATCAGCAGTCAACCAACCGTTATTCTGATGGC
CTGAGCGAGGCAAATACCGCATCGTGTAAAGGACAATTACAAACAGGAATGAACTGCAACGGCGCA
GGAACACTGCCAGCGCATCAACAAATATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCTGTT
TTCCCGGGATCGAGTGGTGGAGTACCATGCAATCAGGAGTACGGATAAAATGCTTGTGAG
AGGCATAAATTCCGTAGCCAGTTAGTCTGACCATCTCATCTGTAACATCATTGCAACCTACCTTGC
CATGTTTCAAAACAACTCTGGCGCATGGGCTCCCATACAATCGATAGATTGTCGCACCTGATTGCCG
ACATTATCCGAGGCCATTACCCATATAAACTCAGCATCCATGTTGAAATTAAATCCGGCCTCGAGCA
AGACGTTTCCGTTGAATATGGCTCATAACACCCCTGTATTACTGTTATGTAAGCAGACAGTTTATTG
TTCATGATGATATAATTATCTTGCAATGAACTCAGAGATTGAGACACAACGGTGGCTTCC
CCCCCCCATTATTGAAGCATTTATCAGGGTATTGCTCATGAGCGGATACATATTGAAATGTTAGAA
AAATAAAACAAATAGGGGTTCCGCGCACATTCCCGAAAGTGCACCTGACGCTAAGAAACCAATTATTA
TCATGACATTAACCTATAAAATAGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:133)

ATGAAGGCCATCTGGTGTGCTGTACACCTTCGCCACCGCCAACGCCGACACCCGTGATCGCTA
CCACGCCAACACAGCAGCACACCGTGGACACCGTGTGGAGAAGAACGTGACCGTGACCCACAGCGTGA
ACCTGCTGGAGGACAAGCACAACGCCAAGCTGCAAGCTCCGGCGTGGCCCCCTGACCTGGCAAG
TGCAACATGCCGCTGGATTCTGGCAACCCCGAGTGTGCGAGAGCCTGAGCACCCGCCAGCAGCTGGAGCTA
CATCGTGGAGACCCCGAGCAGCGAACCGCACCTGCTACCCCGCGACTTCATCGACTACCGAGGACTGC
GGGAGCAGTGGAGCGCTGGAGCAGCTTCGAGGCTGGAGATCTTCCCAAGACAGCAGCTGGCCAAAC
CACGACAGAACAAAGGGCGTGGCCCTGCCCCCACGGCGCAAGAGCTTCTACAAGAACCTGAT
CTGGCTGGTGAAGAAGGGCAACAGCTACCCCAAGCTGAGCAAGAGCTACATCAACGACAAGGGCAAGGAGG
TGCTGGCTGTGGGATCCACCAACCCAGCACCAGCCGAGCAGAGCTGTACCGAGACGGCTGATT
ACCTACGTGTTCTGGCAGCAGCCGGTACAGCAAGAAGTTCAAGCCGAGATGCCATCCGGCCAAGGT
GGGGGACCAAGGGGGGGATGAACACTACTACTGACCCCTGGTGGAGCCGGGACAAGATCACCTTCGAGG
CCACCGGCAACCTGGTGTGCCCCGGTACGCCATGGAGCGGAACGCCGGCAGCGGATCATCATC
AGCGACACCCCGTGCAGACTGCAACACCACCTGCCAGCCCCAAGGGCGCCATCACACCCAGCTGCC
CTTCCAGAACATCCACCCATCAGCATCGCAAGTGGCCAACACTACGTGAAGAGCAGCAAGCTGCC
CCACCGGCCCTGGAAACATCCCGAGCATCCAGAGCCGGGCGTGGTGGCGCCATGCCGGCTTCATCGAG
GGCGGCTGGAGCGGCGATGGTGGAGCGCTGGTACCGCTACCCACCGAGAACGAGCAGGGCAGCGGCTACGC
CGCCGACCTGAAGAGCACCCAGAACGCCATCGACGAGATCACCAACAGGTGAACAGCGTATCGAGAAGA
TGAACACCCAGTTCAACCGCCGTTGGCAAGGGAGTTCAACCACTGGAGAAGCGGATCGAGAACCTGAACAAG
AAGGTGGACGACGGCTTCTGGACATCTGGACCTACAACGCCAGCTGCTGGTGTGCTGGAGAACGAGCG
GACCCCTGGACTACCACCGACAGCAACGTGAAGAACCTGTACAGAGAAGGGTGGAGCCAGCTGAAGAACACG
CCAAGGAGATCGGCAACGGCTGCTCGAGTTCTACCCACAGTGCACAAACACCTGCATGGAGAGCGTGAAG
AACGGCACCTACGACTACCCCAAGTACAGCGAGGAGGCAAGCTGAACCGGGAGGAGATCGACTCCGGAGG
CGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGATGCAGAGCAGCAACCTGTACATGAGCATGA
GCAGCTGGTGTACACCCACAGCCTGGACGCCGGCTGTTCTGACCGCAGCCGAGGAGTAC

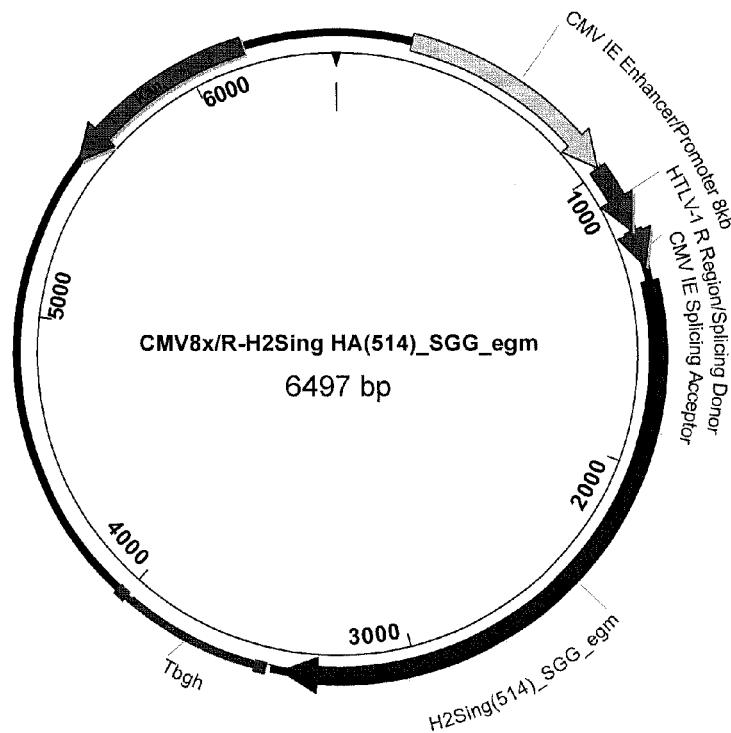
Fig. 26-3

GAGCACGCCAAGAACGCTGATCATCTCCTGAACGAGAACAAACGTGCCCGTGCAGCTGACCAGCATCAGCGC
CCCCGAGCACAAGTCGAGGGCCTGACCCAGATCTCCAGAAGGCCCTACGAGCACGAGCACATCAGCG
AGAGCATCAACAACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCCCTGCAGTGG
TACGTGGCCGAGCAGCACGAGGAGGTGCTGTTCAAGGACATCCTGGACAAAGATCGAGCTGATGCCAA
CGAGAACACGGCCTGTACCTGGCCGACCAAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGATCC

Translation (SEQ ID NO:44)

MKAILVVLLYTFATANADTLICIGYHANNSTDVTVDLKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGK
CNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPN
HDSNKGVTAAACPAGAKSFYKNLILVKKGNSPKLSKSYINDKGEVLVLWGIIHHPSADQQSLYQNA
TYVFVGSSRYSKKFPEIAIRPKVRDQGRMNYYWTLVEPGDKITFEATGNLVPRYAFAMERNAGSGII
SDTPVHDNCNTTCQTPKGAIANTSPLPFQNMTHPIITIGKCPKVVKSTKRLATGLRNIPSTIQSRLFGAIAGFIE
GGWTGMVDGWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVMSVIEKMNTQFTAVGKEFNHLEKRIENLNK
KVDDGFLDIWTYNAELLVLLLENERTLDYHDSNVKNLYEKVRSQQLNNAKEIGNGCFEFYHKCDNTCMESVK
NGTYDYPKYSEEAKLNREEIDSQGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEY
EHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQW
YVAEQHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRKSGS

Fig. 26-4



H2 Sing HA(514)_SGG_egm (H2 1957Sing HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:134)

TCGCGCGTTTGGTGTAGACGGTAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGCTG
TAAGCGGATGCCGGGAGCAGACAAGCCGTCAAGGGCGGTGAGGGGTGTTGGCGGGTGTGGGGCTGGCT
TAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGACCATATGCGGTGTAAATACCGCACAGATGCG
TAAGGGAAAAATAACCGCATCAGATTGGCTATTGGCCATTGCATACGTTGATCCATATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACCGCCATGTCACATTGACTTAGTTATTAGTAATAGTAATCA
ATTACGGGAACCTCCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGGAATTCCAACCTGGC
TGACCGCCCCAACGACCCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGAA
CTTCCATTGACGTCAATGGGTGGAGTTACGGTAAACTGCCACTTGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGGAACTTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCACTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGGTTTG

Fig. 27-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTGTTTGACTCACAAAATCAACGGAACTCCAAAATGCTGAACAACCTCGCCCCATTGA
CGCAAATGGCGGTAGGGTGTACGGTGGAGGTCTATAAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCATAGAACACCGGACCGATCCAGCCTCATCGCT
CGCATCTCTCCCTCACCGCCCGCCCTACCTGAGGCGCCATCCACGCCGGTTGAGTCGCGTTCTGCC
GCCTCCCGCCTGTTGCTCCCTGAACCTGCGTCCGCCGTCTAGTAAGTTAAAGCTCAGGTCGAGACCGG
GCCTTTGTCGCCGCCCTCCCTGGAGGCTACCTAGACTCAGCCGCTCTCCACGCTTGGCTGACCCCTGCTT
GCTCAACTCTAGTTAACCGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTGTGCTGCCGCCGCCACAG
ACATAATAGCTGACAGACTAACAGACTGTTCTTCCATGGCTTTCTGAGTCACCGCTGTCACACG
TGTGATCAGATATCGCGCCGCTAGAGATATCGCCACCATGCGCATCTACCTGATCCTGCTGTTTA
CAGCTGTGCGGGGCGATCAGATCTGTATCGGCTACCACGCCAAATAGCACCAGAGAAGGTGGACACCATC
CTGGAAGAAAATGTGACCGTGACCCAGCCAAAGGATATTCTGGAAAGACCCACAACGCCAAGCTGTGCAA
GCTGAATGGCATTCTCTCTGGAACCTGGCGATTGTTCTATTGCTGGCTGCTGGAAATCTGAGT
GGCATAGACTGCTGCTGCGTACTGAGCTACATCATGGAAAAAGAGAACCCCTAGGGACGGACTGTGT
TACCCCGCAGCTTCACGATTACGAGGAACCTGAAGCACCTGCTGTCCAGCGTGAAGCAGCTCGAGAAAGT
GAAGATCTGCCAACGGATAGATGGACCCAGCATAACAACACAGCGGAAGCAGAGCTGTGCTGTCG
GCAACCCCAGCTTCTCAGAAATATGGCTGGCTGACCAAGAAGGGCTCTAATTATCCTGTCGCCAAGGGC
AGCTACAATAATACAAGCGCGAGCAGATGCTGATTATTGGCGTGCACCCCTAATGATGAGACAGA
GCAGAGAACCCCTGTGACAGAAATGTGGGCACATACGTGCTGTGGGCACCAGCACACTGAATAAGAGAAAGCA
CCCCCGATATTGCCAACAGACCCAAAGTGAATGGACAGGGCGCAGAATGGAAATTCTGGACCCCTGCTG
GATATGTGGGACACCATCAACTTGGAGACGACCCGGAACTGTAATTGCGCCCTGAGTACGGCTCAAGATCAG
CAAGAGAGGCAGCAGCGCATTGAAACAGAGGGCACCTGGAAACTGTGAAACCAACTGTCAGACAC
CTCTGGGCCATTAAATACCACCCCTGCCCTTCCATAATGTCACCCCTGACAAATCGCGAGTGGCTAAG
TACGTGAAGTCTGAGAAACTGGTGTGGGCACAGGACTGAGAAATGTGCCCTAGATCGAGTCAAGGGCCT
GTTGGAGGCATTGCGGGCTTATTGAGGGATGGCAGGGAAATGGTGGATGGTGGTACGGCTATCACC
ACAGCAATGATCAGGGATCTGGCTATGCGCGATAAAAGAGAGCAGGGCTTGAAGGCATCACC
AAACAAAGTGAACAGCTGATCGAGAAGATGAACACCCAGTTGAGGGCTGGAAAGAGTTAGCAATCT
GGAAAGACGGCTGGAAACCTGAACAAGAAAATGGAAGATGGCTCTGGACGTGTGGACATATAATGCCG
AGCTGCTGGTGTGATGAAAACAGAGGACCCCTGGACTTCAAGACAGCAACGTGAAGAACCTGTACGAC
AAAGTGCAGTGCAGCTGAGAGACAATGTGAAAGAGCTGGCAACGGCTGCTTGAGTTCTACCACAAGTG
CGACGACGAGTGCATGAATAGCGTGAAGAACGGCACCTACGACTACCCCTAAGTATGAGGAAGAGAGCAAGC
TGAACAGAAACGAGATCAAGTCCGGAGGCACATCATCAAGCTGTGAAACGAGCAGGTGAAACAGGAGATG
CAGAGCAGCAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACAGCTGGACGGCGGGCTGTT
CCTGTTGACACACGCCGAGGAGTACGAGCACGCCAAGAAGCTGATCATCTCTGTAACGAGAACACAG
TGCCCGTGCAGCTGACCGCATCAGGCCCGGAGCACAAGITCGAGGGCTGACCCAGAATCTCCAGAAC
GCCTACGAGCAGCAGCACATCAGCGAGAGCATCAACAACTCTGGACACGCCATCAAGAGCAAGGA
CCACGCCACCTTCAACTCTGCGTACGTGGTACCTGGCCGAGCAGCACGGAGGAGGTGCTTCAAGGACA
TCTGGACAAGATCGAGCTGATCGGCAACGAGAACACGCCCTGTACCTGGGAGCAGTAGTGAAGGGC
ATCGGCAAGAGCAGGAAGAGGGATCCTAGCATCATCATATTAGTCTGGAAAGGGCAATTGATCCAG
ATCTGCTGTCCTCTAGTGGCAGGATCTGTGTTGCCCCCTGGCTTCCATTGACCCCTGGAAAG
GTGCCACTCCCACCTGCTCTTCTATAAAATGAGGAAATTGCACTGGCATTGCTGAGTAGGTGTCATTCT
ATTCTGGGGGGTGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGAAAGACAATAGCAGGCTGCTGGGA
TCCGGTGGGCTCTATGGTACCCAGGTGTGAAGAATTGACCCGGTTCTCTGGGCCAGAAGAACAG
CACATCCCCCTCTCTGTGACACACCCCTGTCACGCCCTGGTTCTAGTTCCAGCCCCACTCATAGGACAC
TCATAGCTCAGGAGGGCTCCGCTTCAATCCCACCCGCTAAAGTACTGGAGGGCTCTCCCTCCCTCAT
CAGCCCACCAAACCAACCTAGCCTCCAAGACTGGAAGAAATTAAAGCAAGATAGGCTTAAAGTGCAGA
GGGAGAGAAAATGCCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAATTAAAGCCATGTTAA
GGCCATCATGGCTTAATCTTCCGCTTCTCGTCACTGACTCGCTCGCTGGTCGTTGGCTGCCGGA
GCGGTATCAGCTCACTAAAGCGGTAAACGGTATCCACAGAAATCAGGGATAACGCAAGGAAAGAACAT
GTGAGCAAAGGCCAGGAAAGGCCAGGAAACGTAAGGCTAAAGGCCGCTGGCTGGCTTCTCCATAGGCTCC
GCCCGGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCAGAACCCGACAGGACTATAAAGA
TACCAAGGGCTTCTCCCTGGAAAGCTCCCTGTGGCTCTCTGTTCCGACCCCTGCCGTTACCGGATAACCT
GTCCGGCTTCTCCCTGGAAAGCGTGGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGTGT
AGGTGTTGCGTCCAAGCTGGCTGTCAGAACCCCCCGTTCAGCCGACCGCTGCCCTTATCCGGT

Fig. 27-2

AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGAT
 TAGCAGAGCGAGGTATGAGCGGTGCTACAGACTCTGAACTGGTGGCCTAAGCTACGGCTACACTAGAA
 GAACAGTATTGGTATCTGCCTCTGCTGAAGCCAGTTACCTCGAAAAGAGTTGGTAGCTCTGATCC
 GGCAACAAACCACCGCTGGTAGCGGTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAGG
 ATCTCAAGAAGATCCTTGATCTTCTACGGGTCTGACGCTAGTGAACGAAAACTCAGTTAAGGA
 TTTGGTCATGAGATTATCAAAGGATCTCACCTAGATCTTAAATTAAAATGAAGTTAAATCA
 ACTAAAGTATATATGACTAAACTTGGTCTGACAGTTACCAATCTTAATCAGTGAGGCACCTATCTCAGC
 GATCTGTCTATTGTTCATCCATAGTGCCTGACTCGGGGGGGGGGGCGCTGAGGTCTCCCTCGTGAAG
 AAGGTGTTGCTGACTCATACCGGCCIGAATGCCCATCATCCAGCCAGAAAGTGAAGGGAGCCACGGTIG
 ATGAGAGCTTGGTGTAGGCGACCAGTGGTGATTGAACTTTGCTTGCACGGAACGGTCTGGTT
 GTCGGGGAAGATGCGTGATCTGATCCTCAACTCAGAAAAGTCTGATTTATTCAACAAACCGCCGTC
 TCAAGTCAGCAGTAACTGCTGCCAGTGTACAACCAAATTACCAATTCTGATAGAAAAACTCATGAGCA
 TCAATGAAACTGCAATTATTATCATCAGGATTATCAATAACCATATTITGAAAAGCCGTTCTGAT
 GAAGGAGAAAACCTACCGAGGCACTTCCATAGGATGGCAAGATCTGGTATCGGTCTGCAATTCCGACTCG
 TCCAACATCAATACAACCTATTAAATTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCAG
 TGACGACTGAATCCGGTGAAGATGGAAAAGCTTATGCAATTCTTCCAGACTTGTCAACAGGCCAGCA
 TTACGCTCGTCATCAAAACTCGCAGTCAACCAAACCGTTATTCTGATTCGCGCTGAGCAGACG
 AAATACGCGATCGCTGTTAAAGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCA
 CGCGATCAACAATATTTCACCTGAATCAGGATATTCTCTAAACCTGGAACTGCTGTTTCCGGGGATC
 GCAGTGGTGAAGTAACCATGCACTCATCAGGAGTACGGATAAAATGCTGATGGTGGAAAGAGGCAAAATT
 CGTCAGCCAGTTAGTCTGACCATCTCATCTGAACTCATGGCAACGCTACCTTGCCATGTTGAGAA
 ACAACTCTGGCGCATGGGCTTCCATACAATCGATAGATTGTCGACCTGATTGCCGACATTATCGCGA
 GCCCATTTATACCCATATAAAATCAGCATCCATGTTGAATTAAATCGGGCCTCGAGCAAGACGTTCCG
 TTGAATATGGCTCATAACACCCCTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTCATGATGATA
 TATTTTATCTTGCAATGAACTCAGAGATTGAGACACAACGTTGCTTCCCCCCCCCATTAT
 TGAAGCATTATCAGGGTTATTGCTCATGAGCGGACATATTGAAATGTTAGAAAATAACAAAT
 AGGGTTCCCGCACATTCCCGAAAAGTGCACCTGACGTCTAAGAAACCATTATTATCATGACATTAA
 CCTATAAAATAGGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:135)

ATGGCCATCATCTACCTGATCCTGCTTTACAGCTGTGCGGGCGATCAGATCTGTATCGCTACCACGC
 CAACAATAGCACCAGAGAAGGTGGACACCATCTGAAAGAAATGTGACCGTGACCCACGCCAAGGATATT
 TGGAAAGACCCACAACGCAAGCTGTGCAAGCTGAATGGCATTCCCTCTGAACTGGCGATTGTTCT
 ATTGCTGGCTGGCTGCTGGAAATCTGAGTGCAGACTGCTGCTGCTGAGTGGAGCTACATCAT
 GAAAAAGAGAACCCCTAGGGACGGACTGTCTTACCCCGAGCTTCAACGATIACGAGGAACCTGAC
 TGCTCTCCAGCGTGAAGCACTTCAAGAAAAGTGAAGATCTGCCCCAAGGATAGATGGACCCACCA
 ACAGCGGAAGCAGAGCTGTGCTGCTCCGGCAACCCCGAGCTTCAAGAAAATGGTCTGGCTGACCAA
 GAAGGGCTCTAATTATCTGTGCCAACGGCAGCTACAATAACAAGCGGAGCAGATGCTGATTATT
 GGGCGTGCACCACCTAATGATGAGACAGAGCAGAGAACCCCTGACAGGAAATGTGGGCACATACGTCT
 GTGGGCACCAAGCACAATGAAATAAGAGAAGACCCCCGATATTGCCACAGACCCAAAGTGAATGGACAGGG
 CGGCAGAATGGAATTCTGGACCCCTGCTGGATATGTGGGACACCATCAACTTGGAGAGCACCAGGAAATC
 TGATTGCCCCCTGAGTACGCTTCAAGATCAGCAAGAGAGGCAGCAGCGCATCATGAAAACAGAGGGCACC
 CTGGAAAATGTGAAACCAAGTGTCAAGACACCTCTGGCGCCATTAAACACCCCTGCCCTCCATAATGT
 GCACCCCTGTACAATCGCGAGTGCCTTAAGTACGTGAAGTCTGAGAAACTGGTCTGGCACAGGACTGA
 GAAATGTGCCCTAGATCGAGTCAAGAGGCTGTTGGAGCCATTGCCGCTTATTGAAGGGGGATGGCAG
 GGAATGGTGGATGGTGGTACGGCTATCACCACAGCAATGATCAGGATCTGGCTATGCCGCCATAAAGA
 GAGCACCCAGAAGGCCTTGTGAGCGTACCAACAAAGTGAACAGCGCTGATCGAGAAGATGAACACCCAGT
 TTGAGGCCGTGGCAAAAGAGTCTAGCAATCTGAAAGACGGCTGGAAAACCTGAACAAGAAAATGGAAGAT
 GGCTTCTGGACGATATAATGCCGAGCTGCTGGTGTGATGGAAAACAGAGGAGGAGGCCCTGGACTT
 TCACGACAGCAACGTGAAGAACCTGTACGACAAGTGCAGGATGCACTGAGAGAGACAATGTGAAAGAGCTGG
 GCAACGGCTGCTTGTAGTTCTACCAACAGTGCAGACGAGCTGAGTGAAGAACGGCACCTAC
 GACTACCCCTAAGTATGAGGAAGAGAGCAAGTGAACAGAAACGAGATCAAGTCCGGAGGCACATCATCAA
 GCTGCTGAACGGCAGGTGAACAAGGAGATGCAGAGCAGCAACCTGTACATGAGCATGAGCAGCTGGCT
 ACACCCACAGCCTGGACGGCGCCGCTGTTCTGACCAACGCCGAGGAGTACGAGCACGCCAG

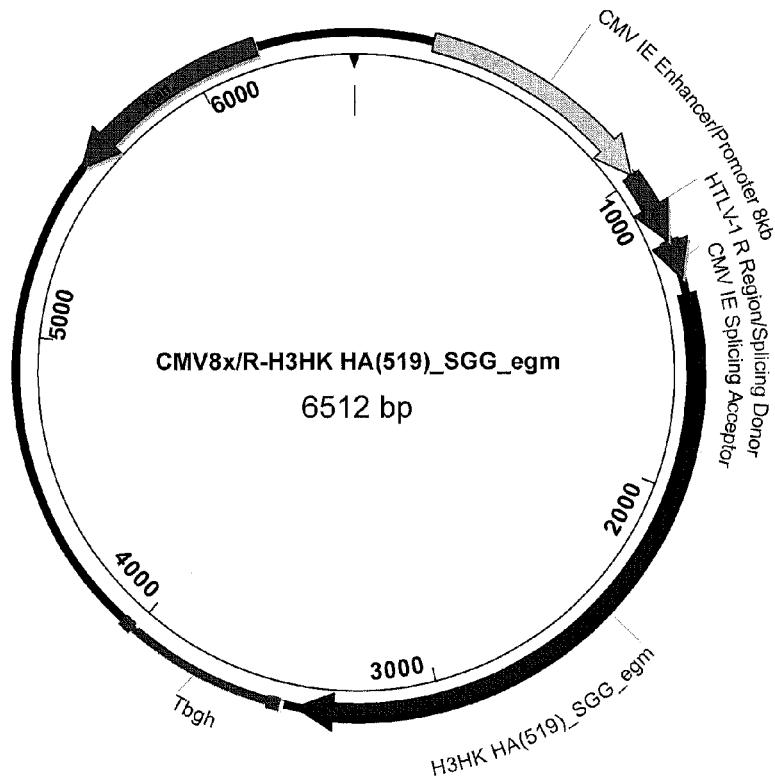
Fig. 27-3

AAGCTGATCATCTCCTGAACGAGAACAAACGTGCCGTGCAGCTGACCAGCATCAGCGCCCCGAGCACAA
GTTCGAGGGCCCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCACATCAGCGAGAGCATCAACA
ACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCCCTGCAGTGGTACGTGCCGAG
CAGCACGAGGAGGGAGGTGCTGTTCAAGGACATCCTGGACAAGATCGAGCTGATCGGCAACGAGAACACGG
CCTGTACCTGGCCGACCAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGATCC

Translation (SEQ ID NO:47)

MAIYLILLFTAVRGDQICIGYHANNSTEKVDТИERNVTVTHAKDILEKTHNGKLCKLNGIPPLELGDCS
TAGWLLGNPECDRLLSVPEWSYIMEKENPRDGLCPGSFNDYEELKHLLSSVKHFEVKILPKDRWTQHTT
TGGSRACAVSGNPSFFRNMWLTKKGSNYPVAKGSYNNTSGEQMLIIWGVHPNDETEQRPLYQNVGTYVS
VGTSTLNKRSTPDIATRPKVNGQGGRMEFSWTLIDMWDTINFESTGNLIAPEYGFKISKRGSSGIMKTEGT
LENCETKCQTPLGAINTTLPFHNVHPLTIGECPKYVKSEKLVLATGLRNVPQIESRGLFGAIAGFIEGGWQ
GMVDGWGYHHNSNDQGSGYAADKESTQKAFDGITNKVNVIEMNTQFEAVGKEFSNLERRLENLNKKMED
GFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNVKELGNGCFEFYHKCDDECMSVKNGTY
DYPKYEEESKLNRNEIKSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEYEHAK
KLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAE
QHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRSKSGS

Fig. 27-4



H3 HK HA(519)_SGG_egm (H3 1968HK HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:136)

```

TCGC CG TTT CCGT GATGACGGT GAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGAGCAGACAAGCCGT CAGGGCGCGTCAGCGGGT GTTGGCGGGT GTCGGGCTGGCT
TAAGTATGCCATCAGAGCTGGCTATTGGCATTGCATACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACTTCCATAGCCCATATATGGAGTTCCCGGTTACATAACTACGGGAATTCCAACCTGGC
TGACCGCCAAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGGAACTTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTG

```

Fig. 28-1

GCAGTACATCAATGGCGTGGATAGCGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTTGTGTTGACTCACAAAATCAACGGAAATTCCAAAATGTCGAACAACCTCGCCCCATTGA
CGCAAATGGCGGTAGCGTGTACGGGGAGGTCTATAAAGCAGAGCTGTTAGTGAAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCCATAGAACGACCCGGACCGATCCAGCCTCCATCGGCT
CGCATCTCTCCCTCACGCCCGCCCTACCTGAGGCCATCCACGCCGTGAGTCGCGTTCTGCC
GCCTCCCGCCTGTGGTGCTCCCTGAACGTGTCGCCGTCTAGGTAAGTTAAAGCTCAGGTGAGACCGG
GCCTTGTCCGGCCTCCCTGGAGCCTACCTAGACTCAGCCGCTCTCCACGCTTGCTGACCGTGCCT
GCTCAACTCTAGTTAACGGGGAGGTGAGCTGAGCAGTACTCGTTGCTGCCGCGCGACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCTTCCATGGGCTTTCTGAGTCACCGTCGTGACACG
TGTGATCAGATATCGGCCGCTTAGAGATATGCCACCATGAAACCATATTGCCCTGAGCTACATCT
TTTGTCTGGCTCTGGCCAGGATCTGCCGCAATGATAATAGCACGCCACCTGTGCTCTGGACACAC
GCCGTGCTTAATGGCACCTGGTAAAACCAATTACCGACGACAGATGAAAGTGACCAATGCCACCGAGCT
GGTGCAGAGCAGCACCGGCAAGATCTGAAACAACCCCCACAGAACCTGGATGGCATCGACTGACCC
TGATCGATGCCCTGCTGGCGATCTCACTGCGACGTGTTCCAGAACGAGACATGGGACTGTTGGAG
AGAAGCAAGGCCCTCAGCACTGCTACCCCTACGATGTGCCGATTACGCCCTCTGAGAACGCTGGC
CAGCAGCGCACACTGAAATTCTACCGAGGGCTTACCTGGACAGCGTGAACGAGAACATGGCGCAGCA
ATGCCGTAAAAGAGGCCCTGGCAGCGCTTCTCAGCAGACTGAACTGGCTGACCAAGTCCGGCAGCACC
TACCCCTGTGCTGAAACGTGACCATGCCAACAACGACAACCTGACAAGCTGTACATCTGGGCGTGACCCA
CCCTAGCACCACAGGAAACAGACCAGCCTGTACGTGAGGCCAGAGCTGACCGTGTCTACAGAC
GGTCCCAGCAGACCATCATCCCCAACATCGAGTCAGAACCTGGGTGCGGCCCTGAGCAGCAGAACATCAGC
ATCTACTGGACCATCGTAAAACCTGGGACGTGCTGGTGTACACAGAACGAACTGATCGACCCAG
AGGCTACTTCAGATCGGACCGGCAAGAGCAGCATCATGAGAACGACGCCACCATGATACTGTATCA
GCGAGTGCATACCCCCAACGGCAGCATCCCCAACGACAAGCCCTTCCAGAACGTAACAAGATCACCTAC
GGGCCCTGCCCTAAGTACGTGAAAGCAGAACACCCCTGAAGCTGGCCACCGGATGAGAAATGTGCCGAGAA
GCAGACAAGAGGCCCTGTTGGGCCATTGCCGCTTATGAGAACGGCTGGAGGGCATGATCGATGGGT
GGTACGGCTTCAGACACCCAGAACCTGAGGGCACAGGACAGGCCGCGATCTGAGTCACACAGGCC
ATCGACCAAGAACGGCAAGACTGAGAAAACCAACGAGAACGAGAACGTTCCACCAAGATCGAGAA
AGAATTCAAGCGAGGTGGAGGGCAGAACCTGGAAAAAAATACGTGGAGGGACACCAAGATCGACTGT
GGAGCTACAATGCCGAACTGCTGGTCGCCCTGGAAAACCAAGCACACCATGACCTGACCGACAGCGAGATG
AATAAGCTGTTGAAAAGACCAAGACGGCAGCTGAGAGAAAACGCCGAGGAATGGCAACGGCTGCTCAA
GATCTACCAACTGCGAACACGCCATCGAGAGCATCAGAACGGCACCTACGACCATGATGTGTACA
GGGACGAGGCCCTGAACACAGATTCCAGATCAAGTCCGGAGGGCAGACATCATCAAGCTGCTGAAACGAG
GTGAACAAGAGGAGATGCGAGCAGCACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACGCC
CGGCCGCCGGCCTGTTCTGACCAAGCCGCCAGGAGTACGAGCACGCCAGAACGCTGATCATCTTCC
TGAACGAGAACACGTCGGCTGAGCTGACCGACATCGGCCAGGAGCATCAGGCCAGGAGCACAAGTGTG
CAGATCTTCCAGAAGGCCCTACGAGCACAGCACATCAGCGAGAGCATCAACACATCGTGGACCAACG
CATCAAGAGCAAGGACCACGCCACCTCAACTTCTGAGTGGTACGTCAGGCCAGCACGAGGAGGAGG
TGCTGTTCAAGGACATCCTGGACAAGATCGAGCTGATCGAACGAGAACGAGGCCCTGTACCTGGCGAC
CACTACGTAAGGGCAGGCCAACAGAGCAGGAAGAGCGGATCCTAGCATCATCATCATTAGTCTGGAAG
GGCGAATTGATCCAGATCTGCTGTGCCCTAGTGTGAGGCCATCTGTTGTTGCCCTCCCCCGTGCCTT
CCTGACCCCTGGAAGGTGCCACTCCACTGTCTTCTTAATAAAATGAGGAAATTGACATCGCATTGCTG
AGTAGGTGTATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGAAGACAATAG
CAGGCATGCTGGGGATGCCGTGGCTCATGGTACCCAGGTGCTGAAGAACATTGACCGGTTCTCTGG
CCAGAAAAGAACAGGCCACATCCCCCTCTGTGACACACCCCTGTCACGCCCTGGTTCTAGTCCAGCC
CCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCCCTCAATCCCACCCGCTAAAGTACTGGAGGG
CTCTCCCTCCCTCATGCCAACAAACCTAGCCTCCAAGAGTGGGAAGAAATTAAAGCAAGATAG
GCTATTAAAGTGCAGAGGGAGAGAAAATGCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAAATT
AAGGCCATGATTTAACGGCATCATGCCCTTAATCTCCGTTCTCGTCACTGACTCGCTGCGCTCG
GTTCCGGCTGCCGAGCGGTATCAGCTCAACTAACGGCGTAATACGGTTATCCACAGAATCAGGGATAA
CGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGTTGCTGGCG
TTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATGACGCTCAAGTCAGGGTGGCGAACCG
ACAGGACTATAAGATACCAAGCGTTCCCTGGAAAGCTCCCTCGTGCCTCCTGACGCC
GCTTACCGGATAACCTGTCCGCCCTTCTCCCTTCCGGAAAGCGTGGCGCTTCTCATAGCTCAGGTAGGT
ATCTCAGTTGGTGTAGGTGCTCGCTCAAGCTGGGCTGTGACGAACCCCGTCAAGCGTGGCGAACCGC

Fig. 28-2

TGCGCCTTATCCGGTAACCTACGTCTGAGTCCAACCGGTAAGACACGACTTATGCCACTGGCAGCAGC
 CACTGGTAACAGCATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTGAAGTGGTGGCTAACT
 ACAGGCTACACTAGAAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCGGAAAAGAGTT
 GGTAGCTCTGATCCGCAAACAAACACCAGCTGGTAGCGGTGTTTTGTTGCAAGCAGCAGATTAC
 GCGCAGAAAAAGGATCTAAGAAGATCCTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAA
 ACTCACGTTAAGGGATTGGTATGAGATTATCAAAAAGGATCTTACACCTAGATCCTTTAAATTAAAAA
 TGAAGTTAAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
 GGCACCTATCTCAGCGATCTGCTATTGTTCATCCATAGTTGCTGACTCGGGGGGGGGGGCGCTGAG
 GTCTGCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCTGAATGCCCATCATCCAGCCAGAAAGTG
 AGGGAGGCCACGGTTGATGAGAGCTTGTGAGGTGGACCAGTTGGTGAATTGAACTTTGCTTGCAC
 GGAACGGCTGCTGGTGGAGATGCGTATGATCCTCAACTCAGCAAAAGTTGATTTATTCAAC
 AAAGCCGCCCTCCGTCAGCTAATGCTCGCAGTGTACAACCAATTAAACCAATTCTGATTAGA
 AAAACTCATCGACATCAAACTGCAATTATTATTCATATCAGGATTATCAATACCATATTTGAAAAA
 AGCCGTTCTGTAATGAGGAGAAAACCTACCGAGGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTC
 TGCAGATTCCGACTCGTCCAACATCAATACAACCTATTAAATTCCCCTGTCAAAATAAGGTTATCAAGTG
 AGAAATCACCATGAGTGAGCAGTGAATCCGGTGAGAATGCCAAAGCTTATGCAATTCCAGACTTGT
 TCAACAGGCCAGCCATTACGCTCGTCAAAACTCGCATCAACCAAAACCGTTATTCACTCGTGAATTG
 CGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATGCAACCGC
 GCAGGAACACTGCCAGCGCATCAACATATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCT
 GTTTTCCCAGGGATCGCAGTGGTGAAGTACCATGCATCATCAGGAGTACGGATAAAATGCTGATGGCG
 AAGAGGCATAAAATTCCGTAGCCAGTTAGTCTGACCCTCTCATCTGTAACATCATGGCAACGCTACCTT
 TGCCATGTTCAAGAACAACTCTGGCGCATGGGCTTCCCATAACATCGATAGATTGTCACCTGATTG
 CCGACATTATCGCGAGCCATTACCCATATAATCAGCATCCATGTTGAATTAAATCGCGGCCCTCGA
 GCAAGACGTTTCCCAGGGATGGCTCATAACACCCCTGTTGATTACTGTTTATGTAAGCAGACAGTTTA
 TTGTTCATGATGATATATTGTTATCTGCAATGTAACATCAGAGATTGAGACACAACGTGGCTTCC
 CCCCCCCCCCATTATTGAAGCATTTACAGGGTATTGTCATGAGCGGATACATATTGAATGTTTTA
 GAAAAATAAACAAATAGGGCTCCGCGCACATTCCCAGAAAGTGCCACCTGACGTCAAGAAACCAATT
 TTATCATGACATTAAACCTATAAAATAGGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:137)

ATGAAAAACCATCATGCCCTGAGCTACATCTTTGCTGGCTCTGGCCAGGATCTCCCGCAATGATAA
 TAGCACCGCACCCTGTGCTGGGACACCACGCCGCTGCTTAATGGCACCTGGTAAAACCATTACCGACG
 ACCAGATCGAAGTGAACATGCCACCGAGCTGGTGCAGAGCAGCAGCACCGCAAGATCTGCAACAAACCC
 CACAGAATCTGGATGGCATCGACTGTACCCCTGATCGATGCCCTGCTGGCGATCCTCACTGCGACGTGTT
 CCAGAACGAGACATGGGACCTGTCCTGGAGAGAACAGCAAGGCTTCAGCAACTGCTACCCCTACGATGTG
 CCGATTACGCCCTCTGAGAACGCCAGAAATGGCGCAGCAATGCCCTGTAAGAGGCCCTGGCAGCGCTTCTCAGCAG
 ACTGAACTGCTGACCAAGTCCGGCAGCACCTACCCCTGTCAGCACTGACCATGCCAACACGACAACT
 TCGACAAGCTGTACATCTGGGCGTGCACCACCCCTAGCACCAATCAGGAACAGACCAGGCTGTACGTGAG
 GCCAGCGCAGACTGACCGTGTCTACAGACGGCTCCAGCAGACCATCATCCCAACATCGAGTCAAGACC
 TTGGGTGCGGCCCTGAGCAGCAGAACATCAGCATCTACTGGACCATCGTGAACCTGGCACGTGCTGGTGA
 TCAACAGCAATGGCAACCTGATCGCCCCCAGAGGCTACTTCAGATGCGGACGGCAAGAGCAGCATCATG
 AGAAGCGACGCCCATCGATACCTGTATCAGCGAGTGCATCACCCCAACGGCAGCATCCCCAACGACAA
 GCCCTTCAAGAACGTGAACAAGATCACCTACGGGCCCTGCCCTAAGTACGTGAAGCAGAACACCTGAGC
 TGGCCACCGCAGGAAATGTGCCAGAACAGCAGAACAGAGGCCCTGGCGCATTGCCGCTTATC
 GAGAACGGCTGGGAGGGCATGATCGATGGTGGTACGGCTTCAGACACCAGAATTCTGAGGGCACAGGACA
 GGCGCCGATCTGAGACTACAGGCGCCATCGACCGAGATCAACGGCAAGCTGAACAGATGTCAGAGA
 AACACCAACGAGAACAGTCCACCGAGATCGAGAAAGAATTCAAGCGAGGTGGAGGGCAGAACCCAGCTGGAA
 AAATACGTGGAGGACACCAAGATCGACCTGTGGAGCTACAATGCCGAACGACTGCTGGTGCCTGGAAAACCA
 GCACACCATCGACCTGACCGACAGCAGAGATGAATAAGCTGTTGAAAAGACCAGACGGCAGCTGAGAGAAA
 ACGCGAGGACATGGGCAACGGCTGTTCAAGATCTACCCACAAGTGCACACGCGCATCGAGAGCATC
 AGAACACGGCACCTACGACCGAGATGTGACAGGGACGGCCCTGAACACAGATTCCAGATCAAGTCGG
 AGGCAGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGATGCAAGAGCAGACCGAACCTGTA
 TGAGCAGCTGGTGCACACCCACGCCCTGGACGGCGCCCTGTTCTGACCGACACGCGCCGAGGAG

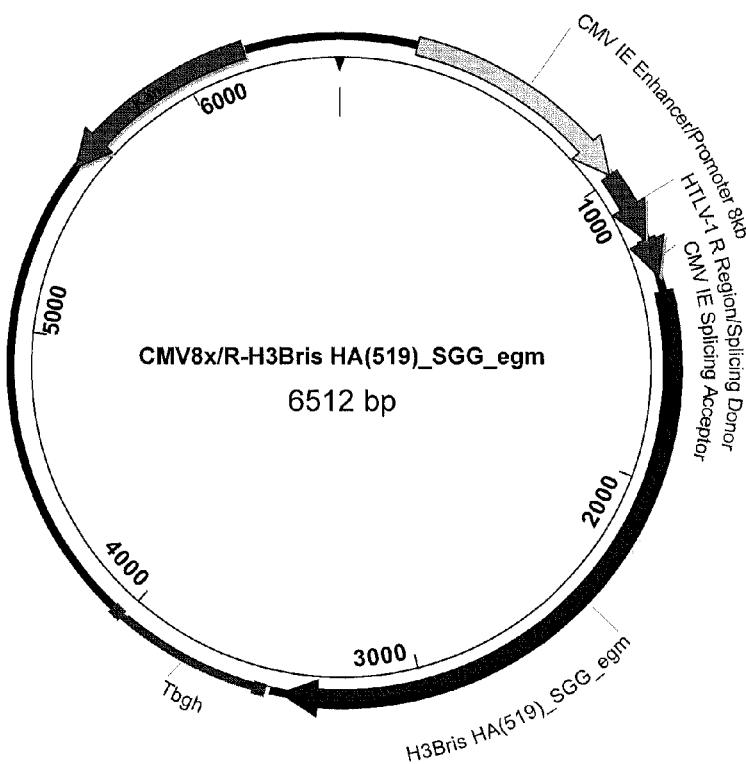
Fig. 28-3

TACGAGCACGCCAAGAACGCTGATCATCTTCCTGAACGAGAACACGTGCCCGTGCAGCTGACCAGCATTCA
CGCCCCCGAGCACAAAGTCGAGGGCCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCACCATCA
GCGAGAGCATCAACAACATCGTGACCACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCCTGCAG
TGGTACGTGGCCGAGCAGCAGGAGGAGGTGCTGTTCAAGGACATCCTGGACAAGATCGAGCTGATCGG
CAACGAGAACACGCCCTGTACCTGGCCGACCAGTACGTGAAGGGATCGCCAAGAGCAGGAAGAGCGGAT
CC

Translation (SEQ ID NO:50)

MKTIIIALSYIFCLALGQDLPGNDNSTATLCLGHAVPNGTLVKTITDDQIEVTNATELVQSSSTGKICNNP
HRILDGIDCTLIDALLGDPHCDVFQNETWDLFVERSKAESNCYPYDVPDYASLRSLVASGTLEFITEGFT
WTGVTQNGGSNACKRGPGSGFFSRLNWLTSGSTYPVLNVTPMPNNDFDKLYIWGVHHPTTNQEQTSLYVQ
ASGRVTVSTRRSQQTIIPNIESRPWVRGLSSRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTGKSSIM
RSDAPIDTCISECITPNGSIIPNDKPFQNVNKITYGACPKYVKQNTLKLATGMRNVPEKQTRGLFGAIAGFI
ENGWEGMIDGWYGFRHQNSECTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
KYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNLFEKTRRQLRENAEDMGNGCFKIHHKCDNACIESI
RNGTYDHDVYRDEALNNRFQIKSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEE
YEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQ
WYVAEQHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRSRSGS

Fig. 28-4



H3 Bris HA(519)_SGG_egm (H3 2007Bris HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:138)

```
TCGC CGCT TCGGT GATGACGGT GAAA ACCTCTGACACATGCAGC TCCCGGAGACGGTCACAGCTTGCTG  
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCTCAGGGGTGTTGGCGGGTGTGGGGCTGGCT  
TAACTATGCCGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGGTGTA  
GAAATACCCACAGATGGC  
TAAGGAGAAAATACCGCATCAGATTGGCTATTGGCATTGCATA  
ACGTTGATCCATATCATAATATGTACA  
TTTATATTGGCTCATGTCCAACATTACCGCATGGTGA  
CATTGATTATTGACTAGTTATTAAATAGTAATCA  
ATTACGGGAACCTCCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGG  
GAATTCCAAACCTGGC  
TGACCGCCCACGACCCCCGCCCATTGACGTCAATAATGACGTATGTT  
CCCATAGTAACGCCAATAGGGAA  
CTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCA  
CTTGGGAATTCCAAGTGTATCATAT  
GCCAAGTACGCCCTATTGACGTCAATGACGGGA  
ACTTCCATAAGCTTGCATTATGCCCA  
GTACATGACC  
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT  
ACCATGGTGTGCGGTTTG
```

Fig. 29-1

GCAGTACATCAATGGCGTGGATAGCGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTGTTGACTCACAAAATCAACGGAAITCCAAAATGTCGTACAACCTCGCCCCATTGA
CGCAAATGGCGGTAGCGTGTACGGTGGGAGGTCTATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACCGCTTTGACCTCCATAGAAGACACCGGACCGATCCACGCCCATCGGCT
CGCATCTCTCCTCACGCCCCGCCACCTGAGGCCATCCACGCCGTGAGTCGCTCTGCC
GCCTCCGCCGTGGTGCCTCCTGAACCTCGCTCCGGCTAGGTAAGTTAACGCTAGGTCGAGACCCG
GCCTTGCGCCTCCCTGGAGCTACCTAGACTCAGCGCTCTCCACGCCCTTGCGCTGACCCCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCGAGCAGTACTCGTGTGCGCGCCACCCAG
ACATAATAGCTGACAGACTAACAGACTGTTCCATGGTCTTCTGAGTCACCGTCGACACG
TGTGATCAGATATCGCGCCGCTAGAGATATGCCACCATGAAACCATATTGCCCTGAGCTACATCC
TGTGCGCTGGTGTACACAGAAGCTGCCGCAACGATAATAGCACGCCACACTGTGCTGGGACACCAC
GCCGTCCTAAAGCACCACGTGAAAACAATCACCAACGACCGAGATCGAAGTGAACATGCCACAGAGCT
GGTGCAGAGCAGCACAGCGAGATCTGTGACAGCCCCCACCAGATCTGGATGGCGAGAACTGTACCC
TGATCGATGCCCTGCTGGCGATCTCAGTGCACGGCTTCCAGAACAGAAATGGGACCTGTTGAG
AGAAGCAAGGCTACAGCACTGCTACCCCTACGACGTGCTGATTACGCCAGCTGAGAACGCTGGG
CTCTAGCGGACCCCTGAAATTCAACACGAGAGCTCAACTGGACCGCGTACACAGAAATGGCACCA
GCCGCTGCATCAGACGGTCCAACACAGCTTCAGTAGACTGAAATTGGCTGACCCACCTGAAGTCAAG
TACCCGCCCTGAACGTGACCATGCCAACATGAGAACGTTGACAAGCTGACATCTGGGAGGTGACCA
CCCTGGCACCGACAACGATCAGATCTTCCCTACGCCAGGCCAGCGCAGAACATACCGTGTCCACCA
GAAGCCAGCAGACCGTGAACCCCAATATCGGCAGCAGACCCAGAGTGCAGAACATCCCCAGCAGGATCAGC
ATCTACTGGACAATCGTGAAGCCTGGCAGACATCTGCTGATCAACAGCACGGCAACCTGATGCCCTCG
GGGCTACTTTAGATCAGAAGCGGAAGAGCAGCATCATGAGATCCGACGCCATCGCAAGTCAAG
GCGAGTGCATCACCCAAACGGCAGCATCCCCAACGACAAGCCCTCCAGAACGTGAACAGGATCACCTAC
GGCGCCTGCCCTAGATACGTGAAGCAGAACACCTCTGAAGCTGGCCACCGGATGAGAAAATGTGCC
GCAGACCAGAGCAGTCTGGCGCCATTGCCGCTTATGAGAACATGGCTGGAGGGAAATGGTGGATGG
GGTACGGCTCAGACACCAGAAATAGCGAGGGAAATTGGACAGGCCGATCTGAAATCTACCCAGGCC
ATCGACCAGATCAACGGCAAGCTGAACAGGCTGATCGGCAAGACCAACGAGAACAGTCCACCA
AGAATTCAAGCAGGGTGGAGGGCAGAATCCAGGACCTGGAAAAAAATACGTGGAGGGACACCAAGATCGAC
GGAGCTACAATGCCAAGTCTGGTCCCTGGAAAACCAGCACACAATTGATCTGACAGACAGTGA
AATAAGCTGTTGAGAAAACCAAGAACAGCAGCTGAGAGAAAACGCCGAGGACATGGCAACGGCTGCTCAA
GATCTACCACAAGTGCACACGCCCTGCATCGGCAGCATCAGAAACGGCACCTACGACCAACGACGTG
GAGATGAGGCCCTGAACAACCGGTTTCAGATCAAGTCCGGAGGGCAGCATCATCAAGCTGCTGA
GTGAACAAGGAGATGCAGAGCAGCAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACGCC
CGGCGCCGGCCCTGTTCTGACGCCAGGCCGAGGAGTACGAGCACGCCAGAACAGCTGATCATCTTCC
TGAACGAGAACACAGTGCACACGCCCTGCAGCTGACCAGCATCAGGCCGAGCACAAGTCCAGGGCTGACC
CAGATCTTCCAGAAGGCTACGAGCAGGAGCAGCACATCAGCAGAGAGCATCAACAAACATCG
CATCAAGAGCAAGGACCACGCCACCTCAACTTCTGCACTGAGCTGATCGGCAACGAGAAC
TGTGTTCAAGGACATCTGGACAAGATCGAGCTGATCGGCAACGAGAACACGCCCTGTACCTGGCGAC
CAGTACGTGAAGGGCAGCAGAGAACAGGAGGAGCAGCACGCCAGGCCAGAACAGCTGATCATCTTCC
GGGAATTGATCCAGATCTGCTGCCCCCTAGTGCCTCTGACACCCCTGCCACGCCCTGGTTCTAGTCC
CCTGACCCCTGGAAGGTGCCACTCCACTGCTCTTCTAATAAAATGAGGAATTGACATCGCATTGTCTG
AGTAGGTGTCATTCTATTCTGG
CAGGCATGCTGGGGATGCGGTGGGCTCTATGGGTACCCAGGTGCTGAAGAACGACCCGGTTCTC
CCAGAAAAGCAGGCACATCCCCCTCTGTGACACCCCTGCCACGCCCTGGTTCTAGTCC
CCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCAATCCCACCCCTAAAGTACTTG
CTCTCCCTCCCTCATGCCACCAACCAAACCTAGCCTCCAAGAGTGGGAAGAAATTAAAGCAAGATAG
GCTATTAAAGTGCAGAGGGAGAGAAAATGCCCAACATGTGAGGAAGTAATGAGAGAACATAG
AAGGCCATGATTAAGGCCATCATGGCTTAATCTCCGCTCTCGCTACTGACTCGCTGCGCTCG
GTTCGGCTGCAGGAGCGGTATCAGCTACTCAAAGGCGTAATACGGTTATCCACAGAAC
CGCAGGAAAGAACATGTGAGCAAAGGCCAGGAAACCGTAAAAGGCCGCTGCTGGCGT
TTTCCATAGGCTGCCCTGACGAGCATCACAAAATGACGCTCAAGTCAAGTCAAGGAGGTGG
ACAGGACTATAAGAACGAGCTCCCGCTGGAGGCTCCCTCGTGCCTCTCGTCTCC
GCTTACCGGATACCTGTCGCCCTTCTCCCTCGGAGGCTGCGCTTCTCATAGCTCACGCTG
ATCTCAGTTCGGTGAGGTGCTCGTCAAGCTGGCTGTCAGAACCCCCCGTT
CAGGCCGACCGC

Fig. 29-2

TGCGCCTTATCCGTAACATCGTCTTGAGTCCAACCCGTAAGACACGACTTATGCCACTGGCAGCAGC
 CACTGGTAACAGGATTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCTGAAGTGGTGGCCTA
 ACGCTACACTAGAAGAACAGTATTGATCTGCCTCTGCAAGCCAGTTACCTTCGAAAAAGAGTT
 GGTAGCTCTGATCCGCAACAAACCACCGCTGGTAGCGGTGGTTTTGCAAGCAGCAGATTAC
 GCGCAGAAAAAAAAGGATCTAAGAAGATCCTTGATCTTCTACGGGCTGACTGAGTGGAAACGAAA
 ACTCACGTTAAGGGATTGGTCATGAGATTATCAAAGGATCTTACCTAGATCCTTAAATTAAAAA
 TGAAGTTAAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
 GGCACCTATCTCAGCAGTCTGCTAATTGCTCATCCAGTGCCTGACTCAGGGGGGGGGGGCGCTGAG
 GTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCGCCTGAATGCCCATCATCCAGCAGAAAGTG
 AGGGAGGCCACGGTTGATGAGAGCTTGTAGGGACAGTTGGTGAATTGAACTTTGCCAC
 GGAACGGTCTCGTTGCGGAAGATGCGTGAATCTGATCCTCAACTCAGCAAAAGTTCGATTATTCAAC
 AAAGCCGCCCTCCGTCAGCTCAGCGTAATGCTCTGCCAGTGTACAACCAATTACCAATTCTGATTAGA
 AAAACTCATCGAGCATCAAATGAAACTGCAATTATTCATATCAGGATTATCAATACCAATTTTGAAAAA
 AGCCGTTCTGTAATGAAGGAGAAAACCAACGGAGCAGTCCATAGGATGCCAAGATCTGGTATCGTC
 TGCATTCCGACTCGTCCAACATCAATACAACCTATTAATTCCCTCGTCAAAATAAGGTTATCAAGTG
 AGAAAATCACCAGTGAAGTGACGACTGAATCCGGTGAGAATGCCAAAGCTTATGCAATTCTCAGACTGT
 TCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACAAACGGTATTCAATTGTTGATTG
 CGCTGAGCAGACGAAATACGCACTCGCTGTTAAAAGGACAATTACAACAGGAATCGAATGCAACGGC
 GCAGGAACACTGCCAGCGCATCAACAAATTTCACCTGAATCAGGATATTCTCTAATACCTGGATGCT
 GTTTTCCCAGGATCGCAGTGGTGAGTAACCATGCAATCAGGAGTACGGATAAAATGCTTGTGGTGG
 AAGAGGCATAATTCCGTCAGCAGTTAGTCTGACCATCTCATCTGAAACATCATTGGCAACGCTACCTT
 TGCCATGTTCAAGAACAACTCTGGCGATCGGGCTTCCCATACAATCGATAGATTGCGCACCTGATG
 CCGACATTATCGCAGGCCATTATACCCATATAAATCAGCATCCATGTTGGAAATTAAATCGGGCTCGA
 GCAAGACGTTCCCGTTGAATATGGCTATAACACCCCCTGTTATGTTATGTAAGCAGACAGTTTA
 TTGTTATGATGATATATTATCTTGTCATGAACTCAGAGATTGAGACACAACGTGGCTTCC
 CCCCCCCCCCCATTATTGAAGCATTATCAGGGTTATTGCTCATGAGCGGATACATATTGAATGTATTAA
 GAAAATAAACAAATAGGGTTCCGCGCACATTCCCGAAAAGTGCACCTGACGTCTAAGAACCATTA
 TTATCATGACATTAAACCTATAAAATAGGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:139)

ATGAAAACCACATTGCCCTGAGCTACATCCTGTCCTGGTGTTCACACAGAAGCTGCCCGCAACGATAA
 TAGCACGCCAACACTGTGCTGGGACACCAGCCGCTGCTTAATGGCACCCTCGTGAACAAACATCCAACG
 ACCAGATCGAAGTGAACATGCCAACAGAGCTGGTGAGAGCAGCAGCACAGGAGATCTGTGACAGCCCC
 CACCAAGATCCTGGATGGGAGAACCTGACCTCTGATCGATGCCCTGCTGGCGATCCTCAGTGGCAGGGCTT
 CCAGAACAAAGAAATGGGACTGTTGGAGAGAAGCAAGGGCTACAGCAACTGCTACCCCTACGACGTGC
 CTGATTACGCCAGCCTGAGAACGCTGGCCTCTAGGGCACCCCTGGAAATTCAACAAAGGAGCTTCAAC
 TGGACCCGGCTGACACAGAACATGGCACCAGCAGCGCTGCATCAGGGTCAACAAACAGCTTCTCAGTAG
 ACTGAAATTGGCTGACCCACCTGAAGTCAAGTACCCGGCCTGAACGTGACCATGCCAACAAATGAGAAAGT
 TCGACAAGCTGACATCTGGGGAGTGACCCACCTGGCACCGACAACGATCAGATCTCCCTTACGCCCCAG
 GCCAGCGGAGAACATCCCGTGTCCACCAAGAGAACCCAGCAGACCGTGAACCTGGCAGCAGACC
 CAGAGTGGGAAACATCCCCAGCAGGATCAGCATCTACTGGACAACTGTAAGCCTGGGAGACATCTGCTGA
 TCAACAGCACCCGAAACCTGATGCCCTCGGGCTACTTTAACATCAGAACGGCAAGAGCACCATCATG
 AGATCCGACCCCCCATCGCAAGTGAACAGGATCACCTACGGCGCTGCCCTAGATACGTGAAGCAGAACACCCCTGAAGC
 GCCCCCTCGAGAACGAGTCCACCAAGATCGAGAACAGAACATTAGCAGGGCTACAGTGAAGCAGAACACCCCTGAAGC
 TGGCCACCCGGCATGAGAAATGTGCCGAGAACGAGACAGAGGCTCTTGGCGCCATTGGGCTTTATC
 GAGAATGGCTGGGAGGGATGGTGGATGGTGGTAGGCCCTCAGACACCAGAACATGAGGGAAATTGGACA
 GGCGCGATCTGAAATCTACCCAGGCCATCGACCGATCAACGGCAAGCTGAACAGGCTGATGGCA
 AGACCAACGAGAACAGTCCACCAAGATCGAGAACAGAACATTAGCAGGGTGGAGGGCAGAACCTGGAA
 AAATACGTGGAGGGACACCAAGATCGACCTGTGGAGCTACAAATGCCAACTGCTGGTCCCTGGAAAACCA
 GCACACAATTGATCTGACAGACAGTGAATAAGCTGTTGAGAAAACCAAGAACGAGCTGAGAGAAA
 ACCCCCAGGACATGGCAACGGCTGCTCAAGATCTACCAAGATGCGACAAACGCCCTGCATGGCAGCATC
 AGAAACGGCACCTACGACCGACGCTGACAGAGATGAGGCGCTGAACAAACGGTTTCAGATCAAGTCCGG
 AGCCGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGATGCAAGAGCAGCAACCTGTAATGAGCA
 'TGAGCAGCTGGTGCTACACCCACAGCGCTGGACGGCGCCGCGCTTCCGTTGACCGACGCCGAGGAG

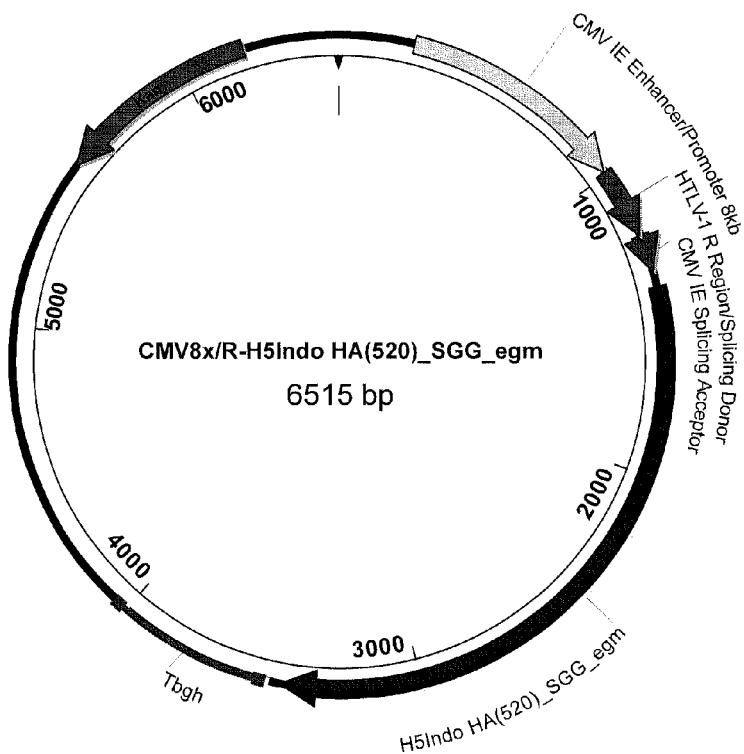
Fig. 29-3

TACGAGCACCCAAGAACGCTGATCATCTCCTGAACGAGAACACGTGCCGTGCAGCTGACCAGCATTCA
CGCCCCCGAGCACAAAGTCGAGGGCTGACCCAGATCTCCAGAAGGCCTACGAGCACGAGCACATCA
GCGAGAGCATCAACAACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTTCAACTTCCATGCAG
TGGTACGTGGCCGAGCAGCACGAGGAGGGTGTGTTCAAGGACATCCTGGACAAGATCGAGCTGATCGG
CAACGAGAACACGGCCTGTACCTGGCCGACCAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGAT
CC

Translation (SEQ ID NO:53)

MKTIIALSYILCLVFTQKLPGNDNSTATLCLGHHAVPNTIVKTITNDQIEVTNATELVQSSSTGEICDSP
HQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERS KAYSNCYPDVPDYASLRSLVASSGTLEFNESFN
WTGVTQNGTSACIRRSNNSFFSRLNWLTHLKFKYPALNVTPNNEKFDKLYIWGVHHPGTDNDQIFPYAQ
ASGRITVSTKRSQQTVIPNIGSRPRVRNIPSRIISIYWTIVKPGDILLNSTGNLIAPRGYFKIRSGKSSIM
RSDAPIGKCNCSECITPNNGSIPNDKPFQNVRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGAIAGFI
ENGWEGMVWDGWYGFRHQNSEGIGQAADLKSTQAAIDQINGKLNRLLGKTNFKHQIEKEFSEVEGRIQDLE
KYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNLFEKTKQLRENAEDMGNGCFKIYHKCDNACIGSI
RNGTYDHDVYRDEALNNRFQIKSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEE
YEHAKKLIIIFLNEENNVPVOLTSISAFEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHTFNFLQ
WVVAEQHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRSRKGSG

Fig. 29-4



H5 Indo HA(520)_SGG_egm (H5 2005Indo HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:140)

```

TCGC CG GTT CCGGT GAT GAC CGGT GAAA ACCT CTG ACAC ATG CAG CTCCGGAGAC CGGT CAC AGCT TGT CTG
TAAG CGG ATGCCGGGAGCAGACAAGCCC GT CAGGGCGCGT CAGCGGGT GT TGGCGGGT GT CGGGGCTGGCT
TAACT ATGCGG CATCAGAGCAGATTGTACTGAGAGTG CACCATATGCGGT GTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGGCCATTGCA TAC GTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATG TCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACTTCCATAGCCC ATATATGGAGTTCCCGC GTTACATAACTACGGGAATT TCCAAACCTGGC
TGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTTGGGAATT TCCAAGTGATCATAT
GCCAAGTACGCCCTATTGACGTCAATGACGGGAACTTCCATAAGCTTG CATTATGCCAGTACATGACC
TTATGGGAATT TCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTG

```

Fig. 30-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTGTTTGACTCACAAAATCAACGGAACTCCAAAATGCTGAACAACCTCCGCCATTGA
CGCAAATGGGGGTAGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCATAGAACAGACCCGGGACCGATCCAGCCTCCATCGGCT
CGCATCTCTCCCTACGCCCGCCCTACCTGAGGCCATCCACGCCGGTTGAGTCGCGTTCTGCC
GCCTCCCGCCTGCGCTCCTGAACCTCCGGCTAGGTAAGTTAAAGCTCAGGTCGAGACCG
GCCTTGTCCGGCGCTCCCTGGAGCCTACCTAGACTCAGCCGGCTCTCACGCTTGCCTGACCCCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCGAGCAGTACTCGTTGCTGCCGCCACCCAG
ACATAATAGCTACAGACTAACAGACTGTTCTTCCATGGGCTTTCTGAGTCACCCTGACACAG
TGTGATCAGATATCGCGCGCTTAGAGATATGCCACCATGGAAAAGATCGTGTGCTGGCCATTG
TGAGCCTGGTGAAGAGCGACCAGATCTGATTGGCTACCAGCACAATAGCACAGAGCAGGTGGACACC
ATCATGGAAAAAAACGTGACCGTGAACGCTCTGATCTGAGAGATTGTAGCGTGGCTGGATGGCTGCTGGCAACCTA
TGTGCGACGAGTTACACGTGCCGAGTGGAGCTATATCGGAGAACGCCAACCCACCAACGATCTG
TGTACCCCGCAGCTTCAACGATTACGAGGAACCTGAAGCACCTGCTGCCGGATCAACCACTTCGAGAA
GATCCAGATCATCCCCAAGTCTCTGGAGCGATCACGAAGCCTTAGCGGAGTGTAGCGCCTGTCCTT
ACCTGGGAGCCCCAGCTTCTCAGAAACGTGGTGTGGCTGATCAAGAACAGCACCTACCCACCATC
AAGAACAGCTACAACAACCCAACCAGGAAGATCTGCTGGCTCTGGGGAAATCCACCAACCTAATGATGC
CCCCGAGCAGACCAGACTGTACCGAGAACCCCCACCCATATCAGCATGCCACAGCACCTGAATCAGA
GACTGGTGCCAAGATCGCACCAGATCCAAGGTGAACGCCAGAGCGGAGGATGGAATTCTGGACCC
ATCCTGAAGCCCAAGCAGCCATCAACCTCGAGAGAACAGGCAACCTTATGCCCTTGAGTACGCCCTACAA
GATCGTGAAGAACGGCGACGCCATCATGAAGAGCGGCTGGAATACGCCAACCTGAACACCAAGTGC
AGACACCTATGGCGCCATCAACAGCAGCATGCCCTTCCACAACATCCACCCCTGACCATCGCGAGTGC
CCTAAGTACGTGAAGAGCACAGACTGGTGTGGCCACAGGCTGAGAAATAGCCCCAGCGGGAGAGCAG
AAGAAAGAGAGGGGCTGGAGGACATCGCCGCTTATTGAGAGCGGCTGGAGGGAAATGGTGGATG
GCTGGTACGGTACCCACACGCAATGAGCAGGCCCTCTGGATATGCCGCCAACAGACTTACCCAGAAG
GCCATCGACGGCTACCAACAAGGTGAACAGCAGCATCGACAAGATGAACACCCAGTTCGAGGCTGTTGG
CAGAGAGTTCAACACACTGGAACGGCGATCGAGAACCTGAACAAGAAAATGGAAGATGGCTCTGGATG
TGTGGACCTACAATGCGAACTGCTGGTGTGATGAAAACGAGCGGACCTGGACTTCCACGACAGAAC
GTGAAGAACCTGTACGACAAGTGCCTGCTGAGAGAACGCAAAGAGCTGGCAACGGCTGCTT
CGAGTTCTACCAACAAGTGCACAACGAGTGCATGAAAAGCATCAGGAACGGCACCTACAACCTACCCCTAGT
ACAGCGAGGAAGCCAGGCTGAAGAGGAAGAGATCAGCTCCGGAGGCGACATCATCAAGCTGCTGAACGAG
CAGGTGAACAAGGAGATCGAGAGCAGCAACCTGTACATGAGCATGACAGCTGGTGTACACCCACAGCCT
GGACGGCGCCGCTGTTCTGACCCAGGCCGCCAGGGACTACGAGCACGCCAACAGCTGATCATCT
TCCGAAACGAGAACACAACGCGCTGAGCTGACCCAGCATCAGCCCCCCCAGGCCAACAGTTCGAGGGCTG
ACCCAGATCTCCAGAAAGCCCTACGAGCACGAGCACATCAGCGAGAGCATCAACAAACATCGTGGACCA
CGCCATCAAGAGCAAGGACACGCCACCTCAACTCTGCACTGGTACGTGGCCAGCAGCACGAGGAGG
AGGTGCTGTTCAAGGACATCTGGACAAGATCGAGCTGATGCCAACGAGAACACCAGGCCCTGACCTGGCC
GACCAACTGCAAGGGCATGCCAAGAGCAGGAAGAGCGGATCTAGCATCATCATCATTAGTCG
AAGGGCAATTGATCCAGATCTGCTGCTCTAGTTGCCACCATCTGTTGTTGCCCTCCCCCTG
CTTCCTGACCTGGAAAGGTGCCACTCCACTGTCCTTCTTAATAAAATGAGGAAATTGATCGCATG
CTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGAAGACAA
TAGCAGGCATGCTGGGATGGTGGCTCTAGGGTACCCAGGTGCTGAAGAAATTGACCCGGTTCTCCT
GGGCCAGAAAGAACGAGGACATCCCTCTGTCGACACACCCCTGTCAGGCCCTGGTCTTAGTCCA
GCCCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCCCTAACATCCACCCGCTAAAGTACTGGAGC
GGTCTCTCCCTCCCTCATGCCAACAAACCTAGCCTCAAGAGTGGGAAGAAATTAAAGCAAGA
TAGGCTATTAGTGCAGAGGGAGAGAAATGCCCTAACATGTGAGGAAGTAATGAGAGAAATCATAGAAT
TTAAGGCCATGTTAACGCCATCATGCCCTTAATCTTCCGCTTCTGCTCACTGACTCGCTGCGCTCG
GTCGTTGGCTGCCGGAGCGTATCGCTCACTAACAGCGGTAAACGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAGGCCGGTGTGG
CGTTTTCCATAGGCTCCGCCCTGACGCCATCACAAAAATGACGCTCAAGTCAGAGGTGGCGAAC
CCGACAGGACTATAAACGATACCGCGTTCCCCCTGGAAGGCTCCCTGTCGCCCTCTCCCTGACCC
GCCGCTTACCGGATACCTGTCGCCCTTCTCCCTGCCGAAGCGTGGCGCTTCTCATAGCTCACGCTGTA
GGTATCTCAGTTGGTGTAGGTGCTCCAGCTGGCTGTGACGAACCCCCCGTTGACGCCGAC

Fig. 30-2

CGCTGCGCCTTATCCGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATGCCACTGGCAGC
AGCCACTGGTAACAGGAATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTGAAGTGGCTGCTA
ACTACGGCTACACTAGAACAGTATTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA
GTTGGTAGCTCTTGATCCGCAAACAAACCACCGCTGGTAGCGGTTTTGTTGCAAGCAGCAGAT
TACCGCAGAAAAAAAGGATCTCAAGAAGATCCTTGATCTTCTACGGGTCTGACGCTCAGTGGACG
AAAACACGTTAAGGGATTGGTATGAGAATTCAAAAGGATCTTACCTAGATCTTTAAATAA
AAATGAAGTTTAAATCAATCTAAAGTATATGAGTAAACTGGTCTGACAGTTACCAATGCTTAAATCAG
TGAGGCACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCTGACTCGGGGGGGGGGGCGCT
GAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCTGAATGCCCATCATCCAGCCAGAAA
GTGAGGGAGCACGGTTGAGAGCTTGTTGAGGTGACAGTTGTTGAGTGGACAGTTGTTGATTGAACTTTGCTTGC
CACCGAACGGTCTGCGTTGCGGAAGATCGTGTGATCTGATCCTCACTCAGCAAAGTTGATTTTATTTC
AACAAAGCCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTACAACCAATTAAACCAATTCTGATT
AGAAAAACTCATCGAGCATCAAATGAAACTGCAATTATTATCATCAGGATTATCAATACCATATTGTA
AAAAGCGTTCTGTAATGAAGGAGAAAACCTACCGAGGCACTTCCATAGGATGGCAAGATCTGGTATCG
GTCTGCGATTCCGACTCGTCAACATCAATACAACCTATTAAATTCCCTCGTCAAAGATAAGGTTATCAA
GTGAGAAATCACCAGTGGACTGAATCCGGTGGAGAATGGCAAAGCTTATGCAATTCTTCCAGACT
TGTCAACAGGCCAGCCATTACGCTGTCATCAAATCACTCGCATCAACCAAACCGTTATTCTGTA
TTGCGCTGAGCGAGACGAAATACCGGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACC
GGCGCAGGAACACTGCCAGCGATCAACAATTTCACCTGAATCAGGATATTCTCTAATACCTGGAAAT
GCTGTTTCCCGGGATCGCAGTGGTAGTGGACCATGTCATCATCAGGAGTACGGATAAAATGCTGATGGT
CGGAAGAGGCATAAAATTCCGTAGCCAGTTAGTCTGACCATCTCATCTGTAACATCATGGCAACGCTAC
CTTGCCATGTTTCAGAAAACAACTCTGGCGATCGGCTTCCCATAACAACTCGATAGAATTGTCGACCTGAT
TGCCCGACATTATCGGAGGCCATTACCCATATAACGATCCATGTTGGAATTAAATCGGGCCT
CGAGCAAGACGTTCCCGTGAATATGGCTCATAACACCCCTGTAACTACTGTTTATGTAAGCAGACAGTT
TTATTGTTCATGATGATATATTGTCATGAACTCAGAGATTGAGACACAACGTT
TCCCCCCCCCCCCATTATTGAAAGCATTATCAGGGTATTGTCATGAGGGACATATAATTGAATGTT
TTAGAAAATAACAAATAGGGGTTCCGCGCACATTCCCGAAAAGTGCCACCTGACGTCTAAGAACCA
TTATTATCATGACATTAAACCTATAAAATAGGCGTATCACGAGGCCATTGTC

Coding sequence (SEQ ID NO:141)

ATGGAAAAGATCGTCTGCTGGCATTGTGAGCCTGGTGAAGAGCGACAGATCTGCAATTGGCTACCA
CGCCAACAATAGCACAGAGCAGGTGGACACCACATGGAAAAAAACGTCAGCGTACCCACGCTCAGGACA
TCTGGAAAAGACCCACAACGGCAAGCTGTGATCTGGACGGCTGAAGCCTCTGATCTGAGAGATTGT
AGCGTGGCTGGATGGCTGCTGGCAACCTATGTGCGACGAGTTCACTCACCGTCCGGAGTGGAGCTATA
CGTGGAGAAGGCCACCCACCAACGATCTGTGTTACCCGGCAGCTCAACGATTACGAGGAACCTGAAGC
ACCTGCTGCCCGATCAACCACTTCGAGAAGATCAGATCATCCCAAGTCTCTTGGAGCGATCACGAA
GCCTCTAGCGGAGTGTCTAGCGCTGCTTACCTGGCAGCCCCAGCTTCTCAGAAACGTTGTTG
GATCAAGAAGAACAGCACCTACCCACCATCAAGAAGAGCTACAACAAACCCAACAGGAAGATCTGCTG
TCCCTGTTGGAAATCCACCCCTAATGATGCCAGCAGCAGACTGTACAGAACCCACCAT
ATCAGCATCGGCCACCAGCACCTGAATCAGAGACTGGTGCCAAGATGCCACCAAGTCAAGGTGAACGG
CCAGAGCGGAGGATGGAATTCTCTGACCATCTGAAGGCCAACGACGCCATCAACTCGAGAGCAACG
GCAACTTTATCGCCCTGAGTACGCTACAAGATCGTGAAGAAGGGCGACAGCGCCATCATGAAGAGCGAG
CTGGAATACGCCACTGCAACACCAAGTCCAGACACCTATGGGCCATCAACAGCAGCATGCCCTTCCA
CAACATCCACCCCTGACCATCGCGAGTGCCCTAAGTACGTGAAGAGCAACAGACTGGTGCTGCCACAG
GCCCTGAGAAATAGCCCCAGCGGGAGAGCAGAAGAAGAAGAGGGCTGTTGGAGCCATGCCGGCTTT
ATTGAAGGCGCTGGCAGGAATGGTGATGGCTGTAACGCTACCAACAGCAATGAGCAGGGCTCTGG
ATATGCCCGACAAAGAGTCTACCCAGAAGGCCATCGACGGCTCACCAACAAGGTGAACAGCATCATCG
ACAAGATGAACACCCAGTTCGAGGCTGTTGGCAGAGAGTTCAACAAACCTGGAACGGCGATCGAGAACCTG
AAACAAGAAAATGGAAGATGGCTTCTGATGTGAGCCTACAATGCCGAACCTGCTGGTGCTGATGAAA
CGAGCGGACCCCTGGACTTCCACGACAGCAACGTGAAGAACCTGTACGACAAGTGCAGCTGAGAG
ACAACGCCAAAGAGCTGGCAACGGCTTCGAGTTCTACCAACAGTGCAGACAACGAGTGCATGGAAC
ATCAGGAACGGCACCTACAACCTACCGCTCAGTACAGCGAGGAAGCCAGGCTGAAGAGGGAGAGATCAGCTC
CGGAGGCGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGAGTGCAGAGCAGCAACCTGTACATGA
GCATGAGCAGCTGGTCTACACCCACAGCCTGGACGGCGCCGCTGTTCTGACCAACGCCGAG

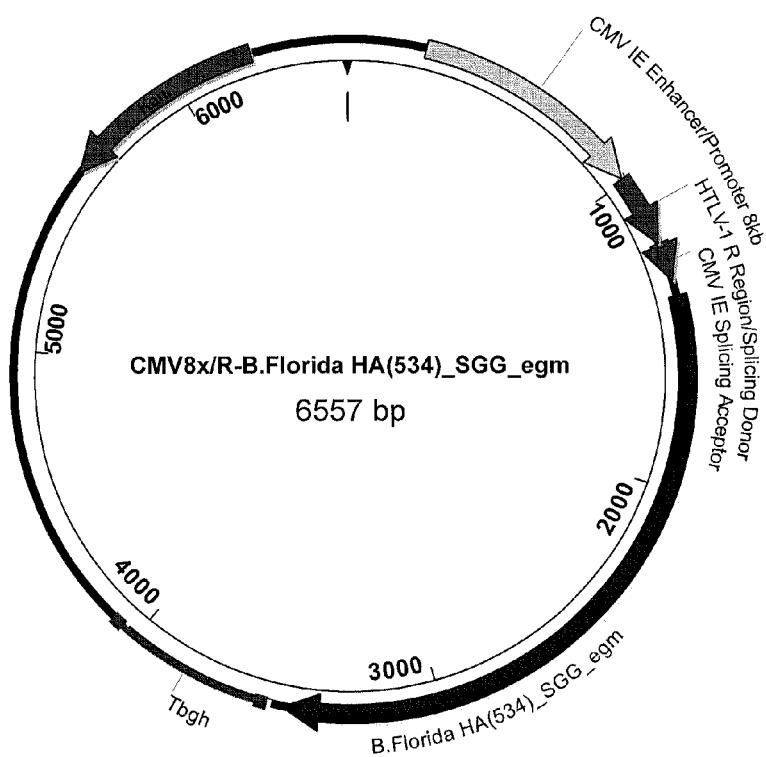
Fig. 30-3

GAGTACGAGCACGCCAAGAAGCTGATCATCTTCCCTGAACGAGAACAAACGTGCCGTGCAGCTGACCAGCAT
CAGCGCCCCCGAGCACAAAGTTGAGGGCCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCAGCAC
TCAGCGAGAGCATCAACAAACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTTCAACTTCTG
CAGTGGTACGTGGCCGAGCAGCACGAGGAGGTGCTGTTCAAGGACATCCTGGACAAGATCGAGCTGAT
CGGCAACGAGAACACGGCCTGTACCTGGCCGACCAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCG
GATCC

Translation (SEQ ID NO:56)

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDLDGVKPLILRDC
SVAGWLLGNPMCDDEFINVPEWSYIIVEKANPTNDLCYPGSFNDYEELKLLSRINHFKEKIQIIPKSSWSDHE
ASSGVSSACPYLGSPSFFRNVVWLKKNSTYPTIKSYNNNTQEDLLVLWGIIHHPNDAAEQTRLYQNPTTY
ISIGTSTLNQRQLVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPHEYAYKIVKKGDSAIMKSE
LEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKSNRVLATGLRNSPQRESRRKKRGLFGAIAGF
TEGGWQGMVDGWGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFBAVGREFNLERRIENL
NKKMEDGFLLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQRLDNAKELGNGCFEFYHKCDNECMES
IRNGTYNPQYSEEARLKREEISSGGDIIKLLNEQVNKEMQSSNLMSMSSWCYTHSLDGAGLFLFDHAAE
EYEHAKKLIIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNPL
QWYVAEQHEEEEVLFKDILDKIELGNENHGLYLADQYVKGIAKSRKSGS

Fig. 30-4



B Florida HA(534)_SGG_egm (B 2006FL HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:142)

TCGCGCGTTTCCGGTGTGATGACGGTGAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGCT
TAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATGCGGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGCCATTGCACTACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACCGCCATGTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACCTCCATAGCCCATAATGGAGTTCCGGTACATAACTAACGGGAATTCCAAACCTGGC
TGACCGCCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTTGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCTATTGACGTCAATGACGGGAACCTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGGTTTG

Fig. 31-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCCATTGACGTCA
ATGGGAGTTGTTGACTCACAAAATCAACGGAACTCCAAAATGTCGTAACAACCTCGCCCCATTGA
CGAAATGGCGGTAGCGTACGGTGGAGGCTATATAAGCAGAGCTGTTAGTGAACCGTCAGATC
GCCGGAGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCAGGACATCCAGCCTCCATCGGCT
CGCATCTCTCCTCACGCCCGCCCTACCTGAGGCCATCCACGCCGGTTGAGTCGCTCTGCC
GCCCTCCGCCGTGGTGCCTCTGAUTGCGTCCGCCGTAGGTAAGTTAAAGCTCAGGTGAGACCGG
GCCTTGTCGGCGCTCCCTGGAGGCTACCTAGACTCAGCGGCTCTCCACGCTTGCTGACCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGTCGCCGCGCCACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCTTCCATGGGCTTTCTGCAGTCACCGTCGACACG
TGTGATCAGATATCGCCGCCGCTAGAGATATGCCACCATGAAGGCCATCATCGTGTGCTGATGGTGG
TGACCCAGCAACGCCGATAGAATCTGACCGGCATCACAGCAGCAATAGCCCCATGTGGTAAAAACAGCC
ACCCAGGGCGAAGTGAATGTGACAGCGTGTACCCCTCTGACCACCACCCACCAAGAGACTACTTCGCAA
CCTGAAGGGCACCAACCAGAGGCAAGCTGTGCCGAGATTGCTGAUTGCACTGCACCGATCTGGATGTGGCTC
TGGGCAGACCTATGTGTGTCGGCACACACCACGCCAAGGCCAGCATCTGCAAGGCCAGTCAGGCTGTG
ACCAGCGGCTGCTTCCCCATCATGCAACGCCAACAGATCAGACAGCTGCCAACCTGCTGAGAGGCTA
CGAGAACATCCGGCTGTCCACCCAGAATGTGATCGATGCCAGAAAGCCCTGGCGGACCTTATAGACTGG
GCACCAAGCCCTCTGTCCCAATGCCACCTCAAGAGCGGCTTTTGCACAATGCCCTGGCCGTGCC
AAGGACAACAAACAAGAACGCCACCAACCCCTGTACCGTGGAGGTGCCCTACATCTGACAGAGGCCAGGA
TCAGATCACAGTGTGGGCTTCCACAGCGACGACAAGACCCAGATGAAGAACCTGTACGGCGACAGCAACC
CCAGAAGTTTACCAAGCGCCAATGGCGTACCCACTACGTGTCCAGATCGCAGCTTCCGAT
CAGACAGAGGATGGCGACTGCCCTAGTCTGGCAGGATCGTGGTGGACTACATGATGCAAGCCTGGCAA
GACCGGCACCATCGTGTACAGAGAGCGTGTGCTGCCCTAGAAAGTGTGGTGTGCCAGCGGCAGGTCTA
AAGTGTCAAGGGCAGCTGCCCTGTGTTGGCGAGGCCACTGTCTGCAAGAAAAGTACGGCGGCTGAAC
AAGAGCAAGCCCTACTACACAGCGAGCACGCCAGGCCATGGCAATTGCCCATCTGGTAAAAACCC
CCTGAAGCTGCCAATGGCACCAAGTACAGACCTCCGCCAGCTGTGAAAGAGAGAGGCTTCTGGCG
CCATTGCCGATTCTGAGCGGCTGGGAGGAATGATTGCCGCTGGCACGGCTACATCTCATGG
GCCCATGGCGTGGCTGGCCGCCGATCTGAAGTCTACCCAGGAAGCCATCAACAAGATCACCAAGAACCT
GAACAGCCTGAGCGAGCTGGAAAGTGAAGAATCTGAGAGACTGAGCGGCCATGGATGAGCTGACAACG
AGATCTGGAAACTGGACGAGAAAGTGGATGATCTCCGCGCCGATACAATTCTCCAGATTGAACTGGCC
GTGCTGCTGTCACAGAGGGCATCATCACAGCGAGGATGAAACACCTGCTGCCCTGGAAAGGAAGCTGAA
GAAGATGCTGGGCCCTCTGCCGTGGAGATCGGCAACGGCTGCTGAGAACAGACAAGTCAACCAGA
CCTGCCCTGGATAGAATGCCGCTGCCACCTTCATGCCGGAGTTGCTGCCCTACCTCGACAGCCTG
AATATCACCTCCGGAGGGCACATCATCAAGCTGTGAACGAGCAGGTGAACAAGGAGATGCAAGCAGCAA
CCTGTACATGAGCATGAGCAGCTGGCTACACCCACAGGCCGAGGCCGAGGCCATCAAGAGCAAGGCCACCT
TCAACTTCTGCCAGTGGTACGTGGCGAGCAGCAGGAGGAGGTGCTGTTCAAGGACATCTGGACAAG
ATCGAGCTGATGCCAACGAGAACACGGCTGTACCTGGCGACCAGTACGTGAAGGGCATGCCAAGAG
CAGGAAGAGCGGATCTAGCATCATCATCATTAGTCTGGAAGGGCAATTGATCCAGATCTGCTGTG
CTTCTAGTTGCCAGCCATCTGTTGCTGCCCTCCCCCTGGCTTCAAGGCTGGAGGTGCTGTTCTGG
ACTGTCCTTCTTAATAAAAGAGGAATTGATCGCATTGCTGAGTAGGTGTCATCTATTCTGGGGGG
TGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGAAGACAATAGCAGGCATGCTGGGATGCCGTGGCT
CTATGGGTACCCAGGTGCTGAAGAATTGACCCGTTCTCTGGGCCAGAAAGAAGCAGGCACATCCCCT
CTCTGTGACACACCCCTGTCCACGCCCTGGTTCTAGTCCAGGCCACTCATAGGACACTCATAGCTCAG
GAGGGCTCCGCCCTCAATCCCACCGCTAAAGTACTTGGAGCGGTCTCCCTCCCTCATGCCACCAA
ACCAAAACCTAGCCTCCAAGAGTGGGAAGAAATTAAAGCAAGATAGGCTATTAAGTGCAGAGGGAGGAGAAAA
TGCCTCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAATTAAAGGCCATGTTAAGGCCATCATGG
CCTTAATCTCCGCTTCCCGCTCACTGACTCGCTGCCCTGGCTTCCGAGGCCGAGCGGTATCAGC
TCACTCAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGCTGCCGTTTTCATAGGCTCCGCCCCCTGAC
GAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
TCCCCCTGGAAAGCTCCCTCGTGCCTCTCGTCCGACCCCTGCCGCTTACCGGATACTGTCCGCCCTTC
TCCCTTCGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGTAGGTCGTTCGC

Fig. 31-2

TCCAAGCTGGGCTGTGACGAACCCCCCGTTAGCCCGACCGCTGCGCCTTATCCGTAACACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
GGTATGTAGGCGGTGCTACAGAGTTGAAAGTGGTGGCCTAACACTGGCTACACTAGAAGAACAGTATTT
GGTATCTGCCTCTGCTGAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCTCTTGATCGGCAAACAAAC
CACCGCTGGTAGCGGTGGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTCTACGGGCTGACGCTCAGTGGAACGAAAACCTCAGTTAAGGGATTTGGTCATG
AGATTATCAAAAAGGAICCTCACCTAGATCCTTTAAATTAAAATGAAGTTTAAATCAATCTAAAGTAT
ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTAT
TTCGTTCATCCATAGTTGCTGACTCGGGGGGGGGCGCTGAGGTCTGCTCGTGAAGAAGGTGTTGCT
GACTCATACCAGGCCGAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTT
GTTGTAGGTGGACAGTTGGTGAACCTTTGCTTTGCCACGGAACGGCTGCGTTGTCGGGAAGAT
GCGTGTCTGATCCCTCAACTCAGCAAAGGTTGATTTTCAACAAAGCCCGTCCCGTCAAGTCAGCG
TAATGCTCTGCCAGTGTACAACCAATTAAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAAC
TGCAATTATTATTCATATCAGGATTATCAATACCATATTGAAAAGCGTTCTGTAATGAAGGAGAAAA
CTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCGACTCGTCCAACATCAA
TACAACCTATTAAATTCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCAGTGAAGTGACGACTGAA
TCCGGTGAGAATGGCAAAGCTTATGCAATTCTTCCAGACTTGTCAACAGGCCAGCATTACGCTCGTC
ATCAAAATCACTCGCATCAACAAACCGTTATTCTGCGTGAATTGCGCTGAGCGAGACGAAATACCGAT
CGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAGCGATCAA
ATATTTTACCTGAATCAGGATATTCTCTAATACCTGGAATGCTTTCCGGGATCGCAGTGGTGAG
TAACCATGCACTCATCAGGAGTACGGATAAAATGCTGATGGTGGAGAGGCAATAATTCCGTCAGCAGT
TTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTGCCATGTTGAGAAACAACCTG
GCATCGGGCTTCCATACAATCGATAGATTGTCGACCTGATTGCCGACATTATCGGAGGCCATTATA
CCCATATAAAATCAGCATCCATGTTGAAATTAAATCGGGCTCGAGCAAGACGTTCCGTTGAATATGGC
TCATAACACCCCTGTATTACTGTTATGTAAGCAGACAGTTTATTGTTCATGATGATATAATTATCT
TGTCAATGTAACATCAGAGATTGAGACACAAACGTGGTTTCCCCCCCCCATTATTGAAGCATT
TCAGGGTTATTGTCATGAGCGGATACATATTGAAATGTTAGAAAAATAACAAATAGGGTCCGC
GCACATTCCCAGAAAGTGCACCTGACGCTAAGAAACATTATTGACATTAACCTATAAAAT
AGCGTATCACCGAGGCCATTGCGT

Coding sequence (SEQ ID NO:143)

ATGAAGGCCATATCGTGTGCTGATGGTGGTGAACAGCAACGCCGATAGAATCTGCACGGCATCACCAG
CAGCAATAGCCCCATGTGGTGAACAGCCACCCAGGGGAAGTGAATGTGACAGGCGTATCCCTCTGA
CCACCACCCACCAAGAGCTACTTCGCCAACCTGAAGGGCACAGAACCCAGAGGCAAGCTGTGCCCGAT
TGCTGAACTGACCGATCTGGATGTGGCTCTGGCAGACCTATGTGTGGCACCCACCATCTGCCAA
GGCAGCATCTGACGAAGTGAAGCTGTGACCAGCGCTGCTTCCCGATCATGACGACCGGACCAAGA
TCAGACAGCTGCCAACCTGCTGAGAGGCTACGAGAACATCCGCTGTCCACCCAGAATGTGATCGATGCC
GAGAAAGCCCTGGCGACCTTATAGACTGGCACAGCGGCTTGTCCCAATGCCACCTCAAGAGCG
CTTTTTGCCACAATGCCCTGGCGTGCCTAAGGACAACAAGAACGCCACCAACCTCTGACCGTGG
AGGTGCCCTACATCTGACAGAGGGGAGGATCAGATCACAGTGTGGGCTCCACAGCGACGACAAGACC
CAGATGAAGAACCTGTACGGCAGACAGCAACCCCCAGAAGTTACCAAGCAGGCCAATGGCGACCA
CTACGTGTCCAGATGGCAGCTTCCGATCAGACAGAGGATGGCGACTGCTCAGTCTGGCAGGATCG
TGGTGGACTACATGATGCAAGACGCTGCAAGACGGCACCATCGTGTATCAGAGAGGCGTGTGCTGCCT
CAGAAAGTGTGGTGTGCCAGCGCAGGTCTAAAGTGTCAAGGGCAGCCTGCTCTGATGGCGAGGCCGA
CTGCTGACGAAAGTACGGCGGCTGAACAAGAGCAAGCCCTACTACACAGCGAGCAGGCCAAGGCCA
TCGGCAATTGCCCATCTGGTGAAACCCCCCTGAAGCTGGCAATGGCACCAAGTACAGACCTCCGCC
AAGCTGCTGAAAGAGAGAGGCTTCTTGGGCCATTGCCGATTCCTGAAAGGCGGTGGAGGGAAATGAT
TGCCGGCTGGCACGGCTACATCTCATGGGGCCCATGGCGTGTGGCGCCGATCTGAAAGTCTACCC
AGGAAGCCATCAACAAGATACCAAGAACCTGAACAGCCTGAGCGAGCTGGAGATGAAAGAATCTGCA
CTGAGCGGCGCATGGATGAGCTGCACACAGAGATCTGGAACTGGACGAGAAAGTGGATGATCTCCGCC
CGATACAATTCCCTCCAGATTGAACTGGCGTGTGCTGCCAACGAGGGCATCATCACAGCGAGGATG
AACACCTGCTGCCCTGGAACGGAAGCTGAAGAAGATGCTGGGGCCCTCTGCCGTGGAGATGGCAACGGC
TGCTCGAGACAAAGCACAAGTGCACACAGACCTGCCCTGGATAGAATGCCGCTGGCACCTCAATGCCGG
CGAGTTCACCGCTGCCATCTGACGCTGAATATCACCTCCGGAGGCGACATCATCAAGCTGCTGAACG

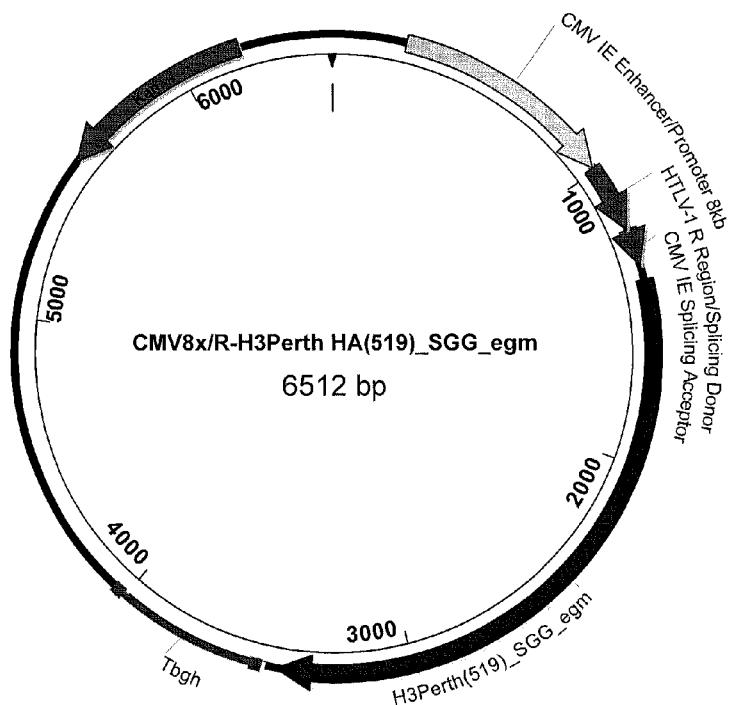
Fig. 31-3

AGCAGGTGAACAAGGAGATGCAGAGCAGAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCCACAGC
CTGGACGGCGCCGGCTGTTCTGTTGACCACGCCCGAGGAAGTACGAGCACGCCAAGAAAGCTGATCAT
CTTCCTGAACGAGAACACACGAGCTGCCGTGCAGCTGACAGCATCAGCGCCCCCGAGCACAAAGTTCGAGGGCC
TGACCCAGATCTTCCAGAAGGCCTACGAGCAGCACATCAGCGAGAGCATCAACAAACATCGGGAC
CACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCCCTGCAGTGGTACGTGGCCGAGCAGCACGAGGA
GGAGGTGCTGTTCAAGGACATCTGGACAAGATCGAGCTGATGGCAACGAGAACCAACGCCGTACCTGG
CCGACCAAGTACGTGAAGGGCATGCCAACAGAGCAGGAAGAGCGGATCC

Translation (SEQ ID NO:59)

MKAIIVLLMVVTNSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPPLTTTPTKSYFANLKGTRTRGKLCPD
CLNCSTDLDVALGRPMCVGTPSAKASILHEVKPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDA
EKAPGGPYRLGTSGSCPNTSKSGFFATMAWAVPKDNNKNATNPNTVEVPYICTEGEDQITVWGFHSDDKT
QMKNLYGDSNPQKFTSSANGVTTHYVSQIGSFDPQTEDGGLPQSGRIVVDYMMQKPGKITGTIVYQRGVILLP
QKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSPPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPA
KLLKERGFFGAIAGFLEGGWEGMIAWGWHGYTSHGAGHVAVAADLKSTQEAINKITKNLINSLELEVKNLQR
LSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNG
CFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSSLNITSGGDIIKLLNEQVNKEMQSSNLMSMSSWCYTHS
LDGACLFLFDHAAEEYEHAKLIIIFLNENNVPQLTISIAPEHKFEGLTQIFQKAYEHEQHISESINNIVD
HAIKSKDHAETNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKSGS

Fig. 31-4



H3 Perth HA(519)_SGG_egm (H3 2009Perth HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:144)

TCGGCGCGTTTGGTGTAGACGGTGAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGACAGACAAGCCCGTCAGGGCGCTCAGCGGGTGTGGCGGGTGTGGGGCTGGCT
TAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGGCATTGCATACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACCGCCATGTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACTTCCATAGCCCATAATATGGAGTTCCGGTGTACATAACTTACGGGAATTCCAAACCTGGC
TGACCGCCCCAACGACCCCGCCCATTGACGTCAATAATGACGTATGTICCCATAGTAACCCCAATAGGGAA
CTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTTGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGGAACTTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGGTTTG

Fig. 32-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTGTTTACTCACAAAATCAACGGGAATTCCAAAATGCTAACAACCTCCGCCATTGA
CGCAAATGGCGGTAGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCATCGGCT
CGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCCATCCACGCCGGTGTAGTCAAGTAAAGCTCAGGTCGAGACCGG
GCCTCCCGCCTGTGGTGCCCTGAACCTGCGTCCGGCTTAGGTAAGTTAAAGCTCAGGTCGAGACCGG
GCCTTGTCCGGCGCTCCCTGGACTCACCTAGACTCAGCCGGCTCTCCACGCTTGCCCTGACCCGCTT
GCTCAACTCTAGTTAACGGTTGGGGCAAGTGTAGTGTAGCGAGTACTCGTTGTCGCCGCGGCCACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCCCTTCATGGGCTTTCTGCACTCACCGTCGACACG
TGTGATCAGATATCGCGGCCGCTAGAGATATGCCACCATGAAAACATAATTGCGCTGTCTACACATAC
TGTGCTGGTTGCCAGAAAACCTGCGGGCAATGACAACCTCAACAGCCACGCTCTGCTTGGGACCCAT
GCCGCCCCAAACGGGACCATGGAAAACCAATTACTAACGATCAGATAGAGGTGACTATGCCACCGAGCT
GGTGCAGGAGCTCCACAGGAGAGATCTGCACTGGGACTTCTGCACTCACCGGAAAGAATTGTACGC
TGATCGACGCGCTGGGGGCAACCTCAGTGTGACGGATTTCAGAATAAGAAGTGGGATCTGTTGAA
AGGTCAAAGGCTTATTCAAATTGCTACCCCTACGATGTCCTGATTATGCCAGCCTGGTCCCTCGTCGC
GTCTAGTGGGACTCTGGAGTTAACACAGAGTCATTAACTGGACTGGGTTACACAGAACGGGACTAGTT
CCGCTTGATAAGGAGAACAAAATAGTTCTTCAGCAGACTGAATTGGCTGACACATCTGAACCTCAAG
TACCCCTGCACTGAATGTAACCATGCCAACACGAGCAGTCCATAACGTTTACATTGGGAGTTCATCA
TCCTGGCACTGACAAGGATCAGATCTTCTGTATGCCAGGCTCCGGCAGGATTACCGTGTCTACAAAGA
GAAGCCAGCAAACCTGTGTCCTCAATATCGGCACTGAGACCCAGAGTACGGAACATCCCTAGTCGCATCAGT
ATTACTGGACCATCGTAAACCAACGGGATATTCTCTGATTAAACAGTACTGGCAACCTGATGCCCGCC
GGGAACTTTAAACCGCTCTGGAAAGTCCATTATGAGATCAGATGCCACCGATCGGAAAGTCAACT
CTGAGTGTATCACACCAATGGGAGCATTCCAAATGACAACACCTTCCAGAACCTTAATCGAATAACCTTAT
GGGGCCTGTCCACGGTACGTGAAGCAAACATCCCTGAAACTGGGACCCGGTATGCGCAATGCCCCAAAA
ACAGACCCGGGGATATTGGGCTATCGCAGGGCTTATCGAGAAATGGCTGGGAAGGGATGGTGGATGGTT
GGTATGGTTTAGACATCAAACCTCCGAAGGCAGAGGCCAGGGCTGCCGATCTCAAGAGCACGCGAGGCGCT
ATAGATCAGATCAATGGAAAGCTAACAGACTGATCGGGAAAACCAACGAAAATTCCATCAGATCGAGAA
AGAGTTCTCGAAGTCGAGGGGCCATACAGGACCTGGAGAAGTATGTTGAGGATACAAAGATTGATCTGT
GGTCTTACAATGCCAGCTGCTGGCTCTGGAGAAATCAGCACACTATTGACCTGACCGATTCAAGAGATG
AACAAACTTTTCAAGAAGACGAAGAACGAGCTTAGAGAAAATCGCAGAGGACATGGGAACGGATGTTAA
AATATATCATAGTGTGATAATGCCCTGCATCGGATCAATTAGAAATGGTACCTATGATCACGATGTTACA
GGGACGAAGCGCTGAATAACAGGTTCCAGATAAAATCCGGAGGCGACATCATCAAGCTGCTGAACGAGCAG
GTGAACAAGGAGATGCCAGAGCAGCAACCTGTACATGACCATGAGCAGCTGGTCTACACCCACAGCTGGA
CGCGCCGGCTGTTCTGTCGACCAACGCCGAGGAGTACGAGCACGCAAGAACGCTGATCATCTTCC
TGAACGAGAACACGTGCCGTGAGCTGACCGACATCAGGCCCGAGGACAAGTTCGAGGGCTGACC
CAGATCTCCAGAAGGCCTACGAGCACGGCAGCACATCAGGAGAGCATCAACACATCGTGGACACGC
CATCAAGAGCAAGGACCAGCCACCTCAACTTCTGCACTGGTACGTGGCAGCAGCACGAGGAGGAGG
TGCTGTTCAAGGACATCTGGACAAGATCAGCTGATCGGCAACGAGAACACCGCTGTACCTGGCGAC
CAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGATCTAGCATCATCATCATATTAGTCTGGAAAG
GGCGAATTGATCCAGATCTGCTGCTTCTAGTTGCCAGCATCTGTTGCCCCTCCCCGTGCTT
CCTTGACCTGGAAGGTGCCACTCCACTGCTCTCTAATAAAATGAGGAAATTGACATCGCATTGTCTG
AGTAGGTGTCAATTCTATTCTGGGGGTGGGGTGGGAGGACAGCAAGGGGAGGATTGGGAAGACAATAG
CAGGCATGCTGGGATGCCGTGGCTATGGTACCCAGGTCTGAAGAATTGACCCGTTCTCTGGG
CCAGAAAAGAAGCAGGCACATCCCTCTGTCGACACACCTGTCCACGCCCTGGTTCTAGTTCCAGCC
CCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCAATCCCACCCGCTAAAGTACTGGAGCGGT
CTCTCCCTCCCTCATGCCACCAAACCAACCTAGCCTCAAGAGTGGGAAGAAATTAAAGCAAGATAG
GCTATTAAAGTGCAGAGGGAGGAGAAAATGCCCTCAACATGAGGAAAGTAAATGAGGAAATCATAGAATT
AAGGCCATGATTAAGGCCATCATGCCCTTAATCTTCCGCTTCTCGCTACTGACTCGCTGCGCTGGTC
GTTCCGGCTGCCGGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAATCAGGGATAA
CGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTGTGGCGT
TTTCCATAGGCTCCGCCCTGCCAGGAGCATCACAAAATGACGCTCAAGTCAAGAGTGGGAACCCG
ACAGGACTATAAGATACCAGGCGTTCCCGCTTCTCCCTCGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGT
ATCTCAGTTCCGCTAGGCTCGCTCAAGCTGGCTGTGACGAACCCCCGTTCAAGCCGACCGC

Fig. 32-2

TGCGCCTTATCCGTAACTATCGCTTGAGTCCAACCCGTAAGACACGACTTATGCCACTGGCAGCAGC
 CACTGGTAACAGGATTAGCAGAGCAGGTATGAGGCGGTCTACAGAGTTCTGAAGTGGGGCTA
 ACAGGCTACACTAGAAGAACAGTATTGGTATCTGCCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTT
 GGTAGCTCTTGATCCGGAAACAAACACCAGCTGGTAGCGGTGGTTTGTGCAAGCAGCAGATTAC
 GCGCAGAAAAAAAGGATCTCAAGAACATCTTGTACATTTCTACGGGTCTGACGCTCAGTGAACGAAA
 ACTCACGTTAAGGGATTGGCATGAGATTATCAAAAAGGATCTCACCTAGATCCTTTAAATTAAAAAA
 TGAAGTTTAAATCAATCTAAAGTATATGAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGA
 GCCACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCCTGACTCGGGGGGGGGCGCTGAG
 GTCTGCCTGTAAGAAGGTGTTGCTGACTCATACCAGGCCTGAATCGCCCCATCATCCAGCAGAAAGTG
 AGGGAGCCACGGTTGATGAGAGCTTGGTAGGTGGACCAGTTGGTGAACCTTGCTTGCAC
 GGAACGGTCTGCGTTGTCGGGAAGATGCGTGTATGATCCTTCAACTCAGCAAAAGTTGATTATTCAAC
 AAAGCCCGTCCCGTCAAGTCAGCCTAATGCTCTGCCAGTGTACAACCAATAACCAATTCTGATTAGA
 AAAACTCATCGAGCATCAAATGAAACTGCAATTATCATATCAGGATTATCAATACCATATTGAAAAA
 AGCCGTTCTGTAATGAAGCAGAAAATCACCGAGGCACTTCCATAGGATGGCAAGATCTGGTATCGGTC
 TCGGATTCCGACTCGTCAACATCAACCTATTAATTCCCTCGTCAAAAATAAGGTTATCAAGTG
 AGAAATCACCAGTGGACTGAGCTGAATCCGGTGGAGAATGGCAAAGCTTATGCAATTCTCAGACTTGT
 TCAACAGGCCAGCCATTACGCTCGTCATAAAATCAGCCTAACCAAAACCGTTATTCAATTCTGATTG
 CGCCTGAGCCAGACGAAATACGGCATCGCTGTTAAAGGACAATTACAAACAGGAAATGCAATGCAACCGGC
 GCAGGAACACTGCCAGCGATCAACAAATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCT
 GTTTTCCCAGGATCGCAGTGTGAGTAACCATGATCATCAGGAGTACGGATAAAATGCTGATGGTCGG
 AAGAGGCATAAAATTCCGTAGCCAGTTAGTCTGACCATCTCATGTAACATCATTGGCAACGCTACCTT
 TGCCATGTTTCAGAAACAACTCTGGCGATCGGGCTTCCATACAATCGATAGATTGTCGACCTGATTG
 CCGACATTATCGGAGCCATTACCCATATAATCAGCATTGTTGGAAATTAAATCCGCCCCCTGA
 GCAAGACGTTCCCGTTGAATATGGCTCATAAACCCCTTGATTACTGTTATGTAAGCAGACAGTTTA
 TTGTTCATGATGATATAATTGCAATGTAACATCAGAGATTGAGACACAACGTGGCTTCC
 CCCCCCCCCCATTATTGAAGCATTATCAGGGTTATTGTCATGAGCGGATACATATTGAATGTTA
 GAAAAATAAAACAAATAGGGTTCGCGCACATTCCCGAAAAGTGCACCTGACGCTAAGAACCATTA
 TTATCATGACATTAACCTAAAAATAGGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:145)

ATGAAAACCATTAATTGCGCTGCTCCTACATACTGTTGCTGGTGTGCCCCAGAAACTGCCGGCAATGACAA
 CTCAACAGCCAGCTCTGCTGGGGCACATGCCGCTTCAACGGGACCATTGTGAAAACCATTACTAACG
 ATCAGATAGAGGTGACTTAATGCCACCGAGCTGGTGAAGTAGCTCCACAGGAGAGATCTGCGATAGTCCC
 CACCAAGATTCTGGACGGAAAGAATIGTACGCTGATCGCGCTGGTGGGACCCCTCAGTGTGACGGATT
 TCAGAAATAAGTGGGATCTGTTGGAAAGGCTTATTCAATTTGCTACCCATTACGATGTGC
 CTGATTATGCCAGCCTCGGCTCGCGTCTAGTGGACTCTGGAGTTCAACAACGAGTCATTAAAC
 TGGACTGGCGTTACACAGAACGGGACTAGTCCGCTTGCATAAGGAGAAGCAAAATAGTTCTCAGCAG
 ACTGAATTGGCTGACACATCTGAACCTCAAGTACCCGACTGTAACATGCCCACAAACGAGCAGT
 TCGATAAGCTTACATTGGGAGTTCATCATCCTGGACTGACAAGGATCAGATCTTGTATGCCAG
 GCTTCCGGCAGGATTACCGTCTACAAAGAGAACGGCAGCAAACGAGTCATTAAAC
 CAGAGTACGGAACATCCCTAGTCGATCAGTATTACTGGACCATCGTGAACCCAGGGCATATTCTCTGA
 TTAACAGTACTGCCAACCTGATGCCCGGGGATACTTTAAATCGCTCGGAAAGTCTCCATTATG
 AGATCAGATGCCGATCGGAAATGCAACTCTGACTGTATCACACCCAAATGGGAGCATTCCAATGACAA
 ACCCTTCCAGAACGTTAATGCAATAACTTATGGGGCTGTCCACGGTACGTGAAGCAAATACCTTGAAC
 TGGCGACGGTATGCGCATGTCGGGAAAGGGATGGTGGATGGTTAGACATCAAACCTCGAAGGCAGAGGCA
 GGCTGCCGATCTCAAGAGCACGCGCCCTAGATCAGATCATGGAAGCTAACAGACTGATCGGG
 AAACCAACGAAAATTCCATCAGATCGAGAAAGAGTTCTCGAACGGCCTACAGGACCTGGAG
 AAGTATGTTGAGGATACAAAGATTGATCTGTTGCTTACAATGCCGAGCTGCTGGCTCTGGAGATCA
 GCACACTATTGACCTGACCGATTAGAGATGAACAAACTTTTGAGAAGAGCAAGCAGCTTAGAGAAA
 ATGCAAGAGGACATGGGAACGGATGCTTAAATATATCATAAAGTGTGATAATGCCCTGACCGATCAATT
 AGAAATGGTACCTATGATCACGATGTTACAGGGACGAAGCGCTGAATAACAGGTTCCAGATAAAATCCGG
 AGGCGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGATGCAAGAGCAGCAACCTGACATGAGCA
 TGAGCAGCTGGTCTACACCCACAGCCTGGACGGCGCCGCTGTTCTGTCGACCGCCGAGGAG

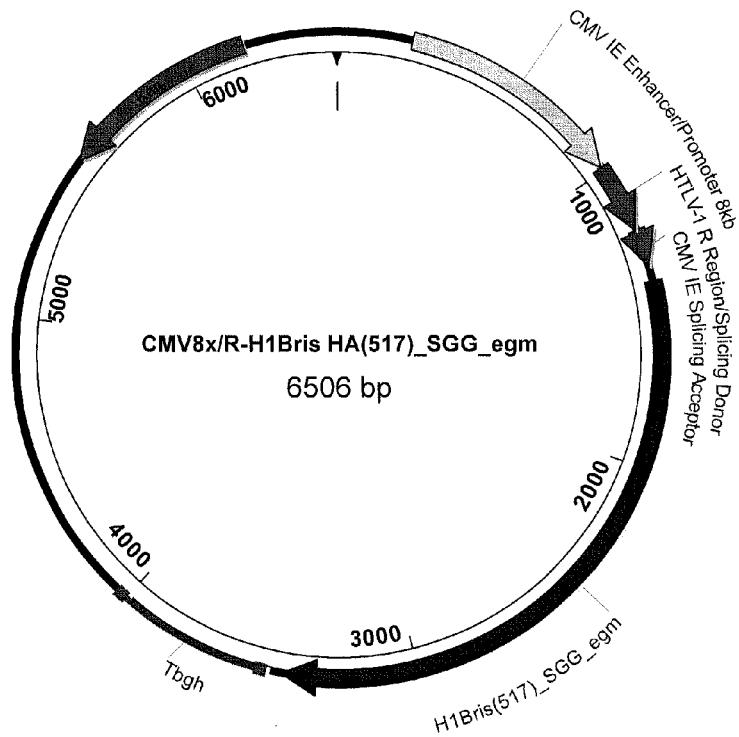
Fig. 32-3

TACGAGCACGCCAAGAAGCTGATCATCTCCTGAACGAGAACACGTGCCGTGCAGCTGACCAGCATCAG
CGCCCCCGAGCACAAAGTTGGAGGGCTGACCCAGATCTTCCAGAAGGCCCTACGAGCACGCCAGCAGCACATCA
GGAGAGCATCAACAAACATCGTGGACCACGCCATCAAGAGCAAGGACCACGCCACCTTCACCTCCTGCAG
TGGTACGTGGCCGAGCACCCAGGAGGTGCTTCAAGGACATCCTGGACAAGATCGAGCTGATCGG
CAACGAGAACACGCCCTGACCTGGCGACCAGTACGTGAAGGGCATCGCCAAGAGCAGGAAGAGCGGAT
CC

Translation (SEQ ID NO:62)

MKTIIALSYILCLVFAQKLPGNDNSTATLCLGHHAVPNTIVKTITNDQIEVTNATELVQSSSTGEICDSP
HQILDGKNCTLIDALLGDPQCDGFQNKKWDLFVERS KAYSNCYPYDVPDYASRLSLVASSGTL FNNESFN
WTGVTQNGTSACIRRSKNSFFSRLNWLT HLNFKYPALNVTPMPNNEQFDKLYIWGVHHPCDKDQIFLYAQ
ASGRITVSTKRSQQTVSPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIM
RSDAPIGKCNECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGAIAGFI
ENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLE
KYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNLKLF EKTKQLRENAEDMGNGCFKIYHKCDNACIGSI
RNGTYDH D VYRDEALNNRFQIKSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEE
YEHAKKLIIFLNENNVPVQLTSIAPEHKFEGLTOIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFQ
WYVAEQHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRSKSGS

Fig. 32-4



H1 Bris HA(517) _ SGG_egm (H1 2007Bris HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:146)

```

TCGCGCGTTTGGTGATGACGGTAAAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGAGCAGACAAGCCGTAGGGCGCGTCAGCGGGTGTGGCGGGGTGTCGGGGCTGGCT
TAACTATGCGGCATCAGAGCAGAATTGACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGGCATTGCACTACGGTGTATCCATATCATAATATGTACA
TTTATATGGCTCATGTCACATTACGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACCTCCATAGCCCATATGGAGTCCCGGTTACATAACCTACGGGAATTCCAACCTGGC
TGACCGCCCAACGACCCCCGGCCATTGACGTCAATTGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATTGGTGGAGTATTACGGTAAACTGCCACTTGGGAATTCCAAGTGTATCATAT
GCCAAGTACGCCCTATTGACGTCAATTGACGGGAACCTCCATAGCTTGCAATTGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTG

```

Fig. 33-1

GCAGTACATCAATGGCGTGGATAGCGTTGACTCACGGAACTTCCAAGTCTCCACCCATTGACGTCA
ATGGGAGTTGTTGACTCACAAATCACGGAAATTCCAAAATGCTGAACAACCTCGCCCCATTGA
CGCAAATGGCGGTAGCGTGTACGGGGAGGCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCGGGACCCATCCAGCTCCATCGCT
CGCATCTCTCCTTCACGCCCGCCGCCCTACCTGAGGCCATCCACGCCGTTGAGTCGCGTCTGCC
CCCTCCCGCTGTGGTGCTCTGAACTCGCTCCGCTTAGGTAAGTTAAAGCTCAGGTCGAGACGG
GCCTTGTCCGGCGTCCCTGGAGCTACCTAGACTCAGCCGCTCTCCACGCTTGCTGACCCGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCCACAG
ACATAATAGCTGACAGACTAACAGACTGTTCTTCCATGGGTTCTTCTGAGTCACCGTCGACACG
TGTGATCAGATATCGCGGCCGCTCTAGAGATATCGCCACCAGTGAAGAGCTGCTGGTGCTGTGTA
CCTTTACCGCCACCTACGCCGATACCATCTGATCGGCTACACGCCAACAAATGACCCGACACCGTGGAT
ACCGTGTGGAAAAGAACGTGACCGTGACCCACAGCGTGAACCTGCTGGAAAACAGCCACACGGCAAGCT
GTGTCTGTAAGGCATTGCCCTCTGAGCTGGAAATTGTAGCGTGCCGGCTGGATTCTGGCAATC
CTGAGTGCAGCTGCTGATTCACAGAGCTCTGCTACATCGTGGAGAAGGCCAACCTGAGAATGGC
ACCTGCTACCCGGCCACTCGCGATTACGAGGAACCTGAGAGAACAGCTGCTCCAGCCTGCTTCGA
GAGATTGGAGATCTTCCCACAGAGAGCAGCTGGCCAATCATACAGTGCACGGCGTGAGCGCCTTGT
GCCCAATGGCGAGAGCAGCTTACAGAACCTGCTGTGGCTGACCCGGCAAGAACGCCACTACAGCA
CTGAGCAAGAGCTACGCCAACAAAGAAAAAGAACGTGCTGGCTCTGGGAGTGCACCCCTCTAA
CATCGGCATCCAGAAGGCCGAGATGCCAAAAGACCCAAAGTGCAGGGGACCCAGGAAGCAGGATCAACTACTGG
GAAAGTTACCCCCGGAGATGCCAAAAGACCCAAAGTGCAGGGGACCCAGGAAGCAGGATCAACTACTGG
ACCCCTGCTGGAACTTGGGACACCATCATCTCGAGGCCAACGGCAATCTGATGCCCTAGATACGCC
TGCCCTGAGCAGAGGCTTGGCAGCGCATTGAGGAGCTATCAATAGCAGCTGCCCTTCCAGAATGTGACGCC
GTCAGACACCACAGGGAGCTATCAATAGCAGCTGCCCTTCCAGAATGTGACCCCTGTGACCATGGCGAG
TGTCCTAAATACGTGCGAGCGCAAGCTGAGAATGGTACCGCCCTGAGGAATATCCCCAGCATCCAGAG
CAGAGGCCATTGGCGCATTGCCGCTTTATCAGGGCGGATGGACAGGCATGGTGGATGGTACG
GCTACCACCAACAGAATGAGCAGGGATCTGGCTATGCCCGATCAGAAGAGCACCCAGAACGCCATCAAC
GGCATCACCAACAAAGTGAACAGCGTATCGAGAAGATGAACACCCAGTTCACCGCCGGGAAAGAGTT
CAACAAGCTGAAACGGGGATGGAAAACCTGAACAAGAAGGTGGACGACGGCTCATCGACATCTGGACCT
ACAACGCCAACCTGGCTCTCTGGAAAATGAGAGGACCCCTGGACTTCCACGACAGAACGTGAAGAAC
CTGTACGAGAAAGTGAAGAGCCAGCTGAAGAACACGCCAACAGAGATCGCAACGGCTGCTCGAGTTCTA
CCACAAGTGCACGACGAGTCAGGAAAGCGTGAAGAACAGCACCTACGACTACCCAAAGTACAGCGAGG
AAAGCAAGCTGAACCGGGAGAAGATCGATTCCGGAGGCACATCATCAAGCTGCTGAACGAGCAGGTGAAC
AAGGAGATGAGCAGAGCAGAACCTGTACATGAGCATGAGCAGCTGTGCTACACCCACGCCGGACGGCG
CCGCCTGTTCTGTCGACCGCCGCCAGGAGTACAGCACGCCAACAGCTGATCATCTCTGAACG
AGAACACAGTGCCTGAGCTGACCATCAGGCCCGAGCACAAGTGCAGGGCTGACCCAGATC
TTCCAGAAGGCATCGACAGCAGCACATCGAGAGACATCAACAAACATCGTGGACCATGCCATCAA
GAGCAAGGACCAAGGCCACCTCAACTCTCGAGCTGAGCTGATGCCAACGAGAACACCAGGCCGTAC
TCAAGGACATCTGGACAAGATCGAGCTGATGCCAACGAGAACACCAGGCCGTAC
GTGAAGGGCATGCCAACAGAGCAGGAAGAGCGGATCTAGCATCATCATCATCATCATCATCAT
TTGATCCAGATCTGCTGTGCTTCTAGTTGCCAGCATCTGTTGCTTGCCTCCCGTGCCTCTGAA
CCCTGGAAGGTGCCACTCCACTGTCCTTCTTAATAAAATGAGGAATTCGATCGCATTGTCTGAGTAGG
TGTCTTCTATCTGGGGGTGGGGCAGGAACAGAACAGGGGGAGGATGGGAAGACAATAGCAGGCA
TGCTGGGGATGCCGCTGGCTCTATGGTACCCAGGCTGAAGAATTGACCCGGTTCTCTCTGGGCCAGAA
AGAACAGGCCACATCCCCCTCTGTGACACACCCCTGTCACGCCCTGGTCTTAGTCCAGCCCCACTC
ATAGGACACTCATAGCTCAGGGCTCCGCCCTAACATCCACCCGCTAAAGTACTTGGAGCGGTCTCC
CTCCCTCATGCCACCAACAAACCTAGCCTCAAGAGTGGAGAAATTAAAGCAAGATAGGCTATT
AAGTGCAGAGGGAGAGAAATGCCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAATTAAAGGCC
ATGATTTAAGGCATCATGCCCTTAATCTTCCGTTCTCGCTCACTGACTCGCTGCGCTGGTCTGCC
CTGCGCGAGCGGTATCGACTCAAAGCGGTAACCGTTATCCACAGAACATCAGGGGATAACGCCAGG
AAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAGGCCGCGTTGCTGGCTTTTCC
ATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCAAACCCGACAGGA
CTATAAAAGATACCAAGCGTTCCCCCTGGAAGCTCCCTGTCGCTCTCCGTGACCGCTTAC
CGGATACCTGTCGCCCTTCTCCCTCGGAAGCGTGGCGTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTAGGTCGTTGCTCCAAGCTGGCTGTGACGAACCCCCCGTTAGCCGACCGCTGCC

Fig. 33-2

TTATCCGTAACATCGTCTTGAGTCCAACCGGTAAAGCACGACTTATGCCACTGGCAGGCCACTGG
 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGGCTAAGTACGGCT
 ACACTAGAAGAACAGTATTTGGTATCTGCCTCTGCTGAAGCCAGTTACCTCGAAAAAGAGTTGGTAGC
 TCTTGATCCGGAAACAAACACCAGCTGGTAGCGGTGTTTTGTTGCAAGCAGCAGATTACGCGCAG
 AAAAAAGGATCTAAGAACAGATCCTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAAACCTAC
 GTTAAGGGATTTGGTCATGAGATTATCAAAAAGATCTCACCTAGATCCTTTAAATTAAAAATGAAGT
 TTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC
 TATCTCAGCGATCTGCTATTCGTTATCCATAGTTCGCTGACTCGGGGGGGGGCGTAGGGTCTGC
 CTCGTGAAGAACGGTGGTCTGACTCATACCAGGCTGAATGCCCATCATCCAGCAGAACAGTGAGGGAG
 CCACGGTGTAGAGAGCTTGTGAGGCTTGAAGTGGACAGTTGGTAGTTGAACCTTTGCTTIGCCACGGAACG
 GTCTGCGTTGCGGAAAGATCGCTGATCTGATCTTCAACTCAGCAAAAGTTCGATTAACTCAACAAAGCC
 GCCGTCGGCTCAAGTCAGCGTAATGCTCTGCCAGTGTACAACCAATTAAACAACTCTGATTAGAAAAGC
 CATCGAGCATCAAATGAAACTCGCAATTATTCAATCAGGATTATCAACCATATTTTGAAGAACGCGT
 TTCTGTAATGAAGGAGAAAACCTACCGAGGAGCTTCCATAGGATGGCAAGATCCTGGTATCGCTGCGAT
 TCCGACTCGTCAACATCAATAACACCTATTAAATTCCCCTCGTCAAAATAAGGTTATCAAGTGAGAAAT
 CACCATGAGTACGACTGAATCCGGTGAAGATGCCAAAAGGTTATGCATTCTCCAGACTTGTCAACA
 GGCCAGGCCATTACGCTCGTATCAAATCACTCGCATCAACAAACCGTTATTCAATTGCGCTG
 AGCGAGACAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGA
 AACACTGCCAGCGCATCAAACATATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCTGTTTC
 CCGGGGATCGCAGTGGTAGTAACCATGCTCATCAGGAGTACGGATAAAATGCTGATGGTGGAGAG
 CATAAATTCCGTCAGCCAGTTAGTCTGACCATCTCATCTGTAACATCATGGCAACGCTACCTTGCAT
 GTTTCAGAAACAACTCTGGCGCATGGGCTTCCACATAATCGATAGATTGCGCACCTGATTGCCGACA
 TTATCGCGAGCCATTATACCCATAAAATCAGCATCCATGTTGAATTAAATCGCGGCTCGAGCAAGA
 CGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTATGTAAGCAGACAGTTTATTGTT
 ATGATGATATATTTATCTTGTGCAATGTAACATCAGAGATTGAGACACAACGTGGCTTCCCCCCCC
 CCCCATTTGAAGCATTATCAGGGTATTGTCATGAGCCGATACATTTGAATGTTAGAAAA
 TAAACAAATAGGGTTCCGCGCACATTCCCGAAAAGTGCCACCTGACGCTAAGAACCACTTATCA
 TGACATTAACCTATAAAATAGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:147)

ATGAAAGTGAAGCTGGTGTGCTGTGACCTTACGCCACCTACGCCATACCATCTGTATCGGCTA
 CCACGCCAACATAGCACCACCGGATACCGTGTGAAAGAACGTGACCGTGAACGCCACAGCGTGA
 ACCTGCTGAAACAGCCACAACGGCAAGCTGTGTGCTGAAAGCATTGCCCTCTGAGCTGGGAAAT
 TGTAGCGTGGCGGCGCTGGATTCTGGCAATCCTGAGTGCAGCTGCTGATTCCAAGAGCTCTGGTCTA
 CATCGTGGAGAACGCCAACCTGAGAATGCCACCTGCTACCCCTGGCAACTTGCAGCTGGGAACTGA
 GAGAACAGCTGCCAGCGTGTCCAGCTCGAGAGATTGAGATCTCCCCAAAGAGAGCAGCTGGGCAAT
 CATACTGACCGCGTGGCGCTTGTAGCCACAATGGGAGAGCAGCTTCTACAGAAACCTGCTGTG
 GCTGACCGCAAGAACGCCCTGTACCCCAACCTGAGCAAGAGCTACGCCAACAAACAAAGAAAAGTGC
 TGGTCCCTGGGAGTGCACCACCCCTCTAACATCGCATCCAGAACGCCCTGTACCAACCCGAGAACGCC
 TACGTGTCCGTGGTGTCCAGCCACTACAGCAGAACGTTACCCCGAGATGCCAAAAGACCCAAAGTGC
 GGACCCAGGAAGGAGCGGATCAACTACTGAGCCCTGCTGGAACTGGCGACACCATCATCTCGAGGCCA
 ACGGCAACTCTGATGCCCTAGATGCCCTTGCCCTGAGCAGAGGCTTGGCAGCGGATCATCAACAGC
 AACGCCCACTGAGAACGCCAGTGTGACGCCAACTGAGCACACCAACAGGGAGCTATCAATAGCAGCTGCC
 CCAGAACATGTCACCCCTGTGACCATCGGAGGTGCTCTAAATACGTGCGGAGGCCAAAGCTGAGAATGGTGA
 CGGGCCTGAGGAATATCCCGACATCCAGAGCAGAGGCGCTGTTGGGCCATTGCGGCTTATCGAGGGC
 GGATGGACAGGCGATGGTGGATGGGTGGTACGGCTACCCACCCAGAATGAGCAGGGGATCTGGCTATGCC
 CGATCAGAACAGAACGCCATCACGGCATCACCAACAAAGTGAACAGCGTATGAGAACAG
 ACACCCAGTTACCCCGTGGCAAAGAGTCAACAGCAGCTGAAACGCCGATGGAAAACCTGAAACAAGAAC
 GTGGACGACGGCTTCATCGACATCTGGACCTACAACGCCGAACCTCTGGTCTCTCTGGAAAATGAGAGGAC
 CCTGGACTTCCACGACAGCAACGTAAGAACCTGTACGAGAACGCTGAAAGAGCCAGCTGAGAACACGCCA
 AAGAGATCGGCAACGGCTGCTCGAGTTCTACCCACAAAGTGAACAGCAGGAGTGCATGGAAAGCGTGAAGAAC
 GGCACCTACGACTACCCCAAGTACAGCGAGGAAAGCAAGCAGCTGAAACCGGAGAAGATCGATTCCGGAGGCCA
 CATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGATGAGCAGAGCAGCAACCTGTACATGAGCATGAGCA
 GCTGGTGTACACCCACAGCCTGGACGGCGCCGCTGTTCTGACCAACGCCGCGAGGAGTACGAG

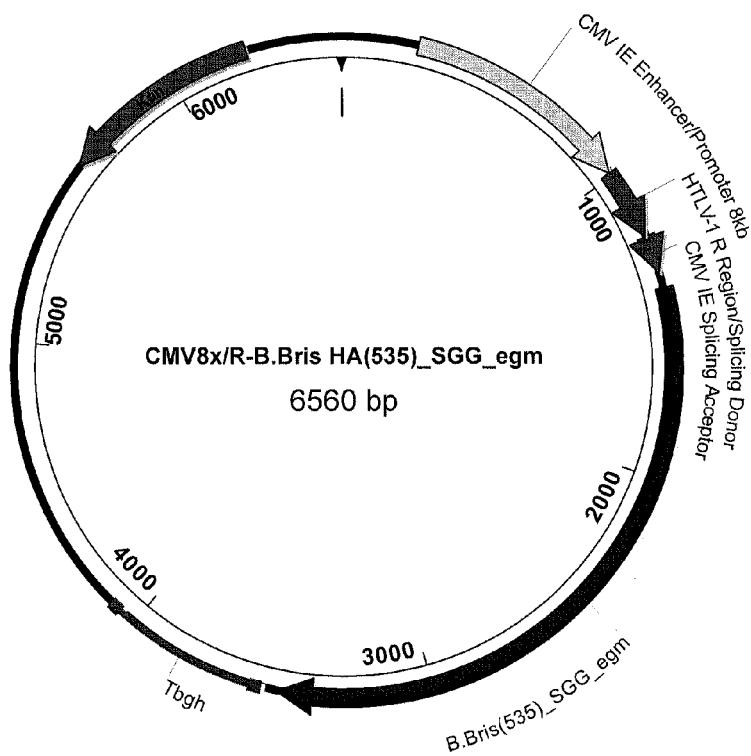
Fig. 33-3

CACGCCAAGAACGCTGATCATCTCCTGAACGAGAACACGTGCCGTGCAGCTGACCAGCATCAGGCC
CGAGCACAAGTCGAGGGCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCACATCAGCGAGA
GCATCAACAAACATCGGGACCACGCCATCAAGAGCAAGGACCACGCCACCTTCAACTTCTGCAGTGGTAC
GTGGCCGAGCAGCACAGGGAGGTGCTTCAAGGACATCTGGACAAGATCGAGCTGATCGGCAACGA
GAACCACGGCCTGTACCTGGCGACCAGTACGTGAAGGGCATCGCCAAGAGCAGGAAGAGCGGATCC

Translation (SEQ ID NO:65)

MKVKLLVLLCTFTATYADTICIGYHANNSTDVTDTVLEKNVTVTHSVNLLENSHNGKLCLLKGTAPLQLGN
CSVAGWILGNPECELLISKEWSYIVEKPNPENGTCYPGHFADYEELREQLSSVSSFERFEIFPKESSWPN
HTVTGVSASCSSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKKEVLVLWGVHHPPNIGIQKALYHTENA
YVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYYWTLLEPGDTIIFEANGNLIAPRYAFALSRGFSGSIINS
NAPMDKCDAKCQTPQGAINSSLPFQNVHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRLFGAIAGFIEG
GWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNNSIEKMNTQFTAVGKEFNKLERRMENLNKK
VDDGFIDIWTYNAELLVLENERTLDFHDNSVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNDECMSVKN
GTYDYPKYSEESKLNREKIDSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEYE
HAKKLIIFLNENNVPVQLTISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHTFNFLQWY
VAEQHEEEVLFKDILDKEI LIGNENHGLYLADQYVKGIAKSRKSGS

Fig. 33-4



B Bris HA(535)_SGG_egm (B 2008Bris HA-ferritin)

Plasmid DNA sequence (SEQ ID NO:148)

TCGCGCGTTCTGGTATGACGGTAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGCGTCAGCGGGTGTGGGGGTGTCGGGGCTGGCT
TAACTATGCCATCAGAGCAGATTGACTGAGAGTGCACCATATGCCGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGCCATTGCATACGTTGTATCCATATCATAATATGTACA
TTTATATTGGCTCATGCCAACATTACGCCATTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACTCCATAGCCATATATGGACTTCCGCGTTACATAACTTACGGGAATTCCAACCTGGC
TGACCGCCCCACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTTGGAAATTCCAAGTGTATCATAT
GCCAAGTACGCCCTATTGACGTCAATGACGGGAACCTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTATGCGGTTTG

Fig. 34-1

GCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGAACTTCCAAGTCTCCACCCCCATTGACGTCA
ATGGGAGTTTGTGTTGACTCACAAAAATCAACGGGAATTCCAAAATGTCGTAACAACCTCCGCCCATGGA
CGCAAATGGCGGTAGGCGTGTACGGTGGAGGTCTATAAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCCTGGAGACGCCATCCACGCTGTTGACCTCATAGAACAGACCCGGGACCGATCCAGCCTCCATCGGCT
CGCATCTCCTTCACCGCCCCGGCCCTACCTGAGGCCGCGATCCACGCCGGTTGAGTCGCGTCTGCC
GCCCTCCGCCGTGGTGCCTCTGAACCTGCGTCCGGCTAGGTAAGTTAAAGCTCAGGTCGAGACCGG
GCCCTTGCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCACGCTTGTGCTGACCCGTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCACTGACTGAGCAGTACTCCTGCTGCCGCGCCACCAG
ACATAATAGCTGACAGACTAACAGACTGTCCTTCCATGGGCTTTCTGAGTCACCGTCGACACAG
TGTGATCAGATATCGGGCCGCTAGATACTGCCACCATGAAGGGCATCATGCTGCTGATGGTGG
TCACAAGCAACGCCGATAGAACTGTAACGGCATCACAGCAGCAATAGCCCTCACGCTGTGAAAACAGCT
ACACAGGGCAAGTGAATGTGACGGCTGATCCCTCTGACCCACAACACTACAAAGAGCCACTCGCCAA
TCTGAAGGGCACAGAGACAAGAGCAAGCTGTCCTCAAGTGCCTGAATTGACAGATCTGGATGTGGCTC
TGGGCAGACCTAAGTGTACAGGCAAATCCCTAGGCGCAAGAGTGTCCATTCTGCATGAAGTGCACCTGTG
ACCAGCGCTGTTTCTATTATGCAACGGCAAGATCACAGACAGCTGCCATTCTGCTGAGAGGCTA
CGAGCACATCAGACTGAGCACCCAAATGTCATCACGCCAAATGCTCTGGGCTGACCAACACTATGTC
GCACATCTGGCAGCTGCCCAACATTACAAATGGCAATGGCTCTTGCACCATGGCTGGGGCGTGCCT
AAGAACGATAAGAACAAAGACGCCACCAACCCCTGACATCGAGGTCCTATATCTGTACAGAGGGCGA
GGATCAGATCACCGTGTGGGAAITTCACAGGCAACAGAACAGATGGCAAGCTGTACGGCGTAGC
AGCCTCAGAAGTTAACGCTGCCAAATGGCTGACCAACACTATGTCAGATCGCCGCTTCCCT
AATCAGACAGAGATGGCGGACTGCCCCTAGCTGGAAGAAATGTTGTTGAAATACATGGTCAGAAGTCTGG
CAAGACGGCAGCATCACATATCAGAGAGGAATCTGCTCCCAAGAGTGTGCGCTTCTGGAAGAT
CCAAAGTGTCAAGGGCAGCTGCCCTGTGATTGGAGAACCGATGTTGTCAGCAGAGAAATACGGCGCCTG
AACAAAGAGCAAGCTTACTATACAGGCGAGCACGCCATCGGCAATTGCTTATTGGGTCAGAC
CCCTCTGAAGCTGGCCAATGGCAAAAGTATAGACCTGCCAGCTGCTGAAAGAGAGAGGCTTTTG
GAGCTATGCCGGCTTCTGGAAGGGCGATGGGAGGGAAATGATTGTCGATGGCATGGCTACACATCTCAT
GGCGCACATGGCGTGGCAGTGGCTGCTGATCTGAAATCTACACAGGAAACCCATACAAGATC
CTGAACAGCCTGAGGGAGCTGGAAAGTGAAGAAATCTGAGACTGTCGAGACTGTCGAG
ATGAGATCTGGAACTGGACGAGAAGGTGGACGATCTGAGAGCCGATACAATCAGCAGCCAGATGAACTG
GCTGTGCTGTCATACGAGGGCATCATCAATAGCGAGGACAACTCTGCTGCCCTGGAAGAAAGCT
GAAGAAGATGTCGGACCTAGGCCGTGGAAATCGGCAATGGATGTTGAGACAAAGCAAGTGAACC
AGACCTGCTGGATAAAATTGCCCCGGAACATTGATGCGGCCGAGTTTCTGCCACCTTCGATAGC
CTGAATATCACATCCGGAGGCGACATCATCAAGTGTGAAACGAGCAGGTGAACAAGGAGATGCGAG
CAACCTGTACATGAGCATGAGCAGTGGTGTACCAACAGCCTGGACGGCGCCCTGTCCTGTT
ACCACGCCGCGAGGGAGTACGAGCACGCCAAAGCTGATCATCTTCTGAAACGAGAACACTGCCGTG
CAGCTGACCGACATCAGGCCCGGAGCAACTGCGAGGGCTGACCCAGATCTTCAGAAGGCCATCGA
GCACGAGCAGCACATCAGCGAGAGCATCAACACATCGGGACCACGCCATCAAGAGCAAGGACACGCCA
CTTCAACTCTGCACTGGTACCTGCGCAGCAGGAGGGAGGTGCTGTTCAAGGACATCTGGAC
AAGATCGAGCTGATGGCAACGAGAACACGGCTGTACCTGGCCGACCAAGTACGTGAAGGGCATGCCAA
GAGCAGGAAGAGCGGATCTAGCATCATCAATCAATTAGTCGGAAGGGCGAATTGATCCAGATCTGCTG
TGCCTCTAGTTGCCAGCATCTGTTGCCCCCTCCCCGTCCTCCCTGACCTGGAAGGTGCCACT
CCCACGTGCTTCTTCTATAAAATGAGGAATTGCAATGCTGAGTAGGTGTCATTCTATTCTGGG
GGGTGGGGTGGGGCAGGACAGCAAGGGGAGGAAATTGGGAAGACATAGCAGGCGATCTGGGATGCCGTGG
GCTCTATGGTACCCAGGTGTCAGAAATTGACCCGGTCTCTGTCGAGGCGAACAGGACATCTGCC
CTTCTCTGTGACACACCTGTCACGCCCTGGTCTGTTGAGTCCAGCCCCACTCATAGGACACTCATAGCT
CAGGAGGGCTCCGCCCTCAATCCACCCGTAAGACTTGGAGCGGTCTCTCCCTCATCAGGCCAC
CAAACCAAACCTAGCTCCAAGAGTGGGAAGAAATTAGCAAGAGTACGGCTATTAGTCAGAGGGAGAGA
AAATGCCCAACATGTGAGGAAGTAATGAGGAATCATAGAATTGACCCGATATTAGGCGCATCA
TGGCTTAATCTCCGCTTCTGCTCACTGACTCGCTGGCTCGTCTGCGCTGCCGAGCGGTATC
AGCTCACTCAAAGGGGTAATACGGTTATCCACAGAAATCAGGGATAACGCAAGGAAAGAACATGAGCAA
AAGGCCAGAAAAGGCCAGGAACCGTAAAGGGCGCTGCTGGCTTTTCTCATAGGCTCCCCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGACTATAAGAAC
GTTTCCCCCTGGAGACTCCCTGCGCTCTGTCGACCCCTGCCGTTACGGGATACTGTCGCC
TTCTCCCTTGGAGACGCGTGGCGTTCTCATAGTCACGCTGTAGGTATCTCAGTTGCGTGAAGTCGTT

Fig. 34-2

CGCTCCAAGCTGGGCTGTGCACGAACCCCCGTTCAAGCCGACCGCTGCCCTTACCGGTAACATCG
 TCTTGAGTCCAACCGGTAAGACACGACTTACGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
 CGAGGTATGAGCGGTCTACAGAGTTCTGAAGTGGTGGCTAACTACGGCTACAGAAGAACAGTA
 TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTGATCCGGCAAACA
 AACACCAGCTGGTAGCGGTGTTTTGTTGCAAGCAGCAGATTACGCCAGAAAAAGGATCTCAAG
 AAGATCCTTGATCTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAACGGGATTTGGTC
 ATGAGATTATCAAAAAGGATCTCACCTAGATCCTTAAATTAAATGAAGTTAACATCTAAAG
 TATATATGAGTAAACTGGTCTGACAGTTACCAATGCTTAATCATGAGGCCCTATCTCAGCGATCTGTC
 TATTTCTGCTATCCATAGTGCCTGACTGGGGGGGGGGGGCGTGAGGCTGCTCGTGAAGAACGGTGT
 GCTGACTCATACCGGGCTGAATGCCCATCATCCAGCCAGAACAGTGGGAGGCCACGGTGTGAGAG
 TTGTTGAGTGGGACCAGTTGGTGAACCTTGCACCGAACGGTCTGCGTTGCGGAA
 GATGCGTGTGATGATCCTCAACTCAGCAAAAGTCGATTATTCAACAAAGCCGCCCTCCGTCAAGTCA
 GCGTAATGCTCTGCCAGTGTACAACCAATTACCAATTCTGATTAGAAAACATCGAGCATCAAATGA
 AACTGCAATTATTCAATACAGATTATCAACCATATTGGAAAAAGCCGTTCTGTAATGAAGGAGA
 AAAACTCACCAGGGCAGTCCATAGGAATGGCAAGATCCTGGTATCGGCTGCAGTCCGACTCGCCAACAT
 CAATACAACTATTAAATTCCCCTCGTCAAAAATAAGTTATCAAGTGAGAACATCACCAGAGTGACGACT
 GAATCCGGTGAAGATGGCAAAAGCTTATGCAATTCTTCCAGACTTGTCAACAGGCCAGCATTACGCTC
 GTCATCAAATCACTCGCATCAACCAACCGTTATCATTCTGATTGCCCTGAGCGAGACGAAATACGC
 GATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCCGCCAGAACACTGCCAGCGCATCA
 ACAATATTTCACCTGATCAGGATATTCTCTAAACACCTGGATGCTGTTTCCGGGATCGCAGTGGT
 GAGTAACCATGCACTCAGGATCAGGATAAAATGCTGATGGTCCGAAGAGGCAATAATTCCGTCAAGCC
 AGTTTAGTCTGACCATCTCATCTGTAACATCAGCTGGCAACGCTTGTCCATGTTTCAAGAACAAACTCT
 GGCGCATCGGCTTCCCATACAATCGATAGATTCTGCACCTGATTGCCGACATTATCGCGAGCCCATT
 ATACCCATATAATCAGCATCCATGTTGAATTAAATCGCGGCCCTGAGCAAGACGTTTCCGTTGAATAT
 GGCTCATAACACCCCTTGTATTACTGTTATGTAAGCAGACAGTTTATTGTCATGATGATATAATT
 TCTTGTCGAATGTAACATCAGAGATTGAGACACAACGTTGCTTCCCCCCCCCATTATTGAAGCAT
 TTATCAGGGTATTGTCATGAGCGGATACATATTGAAATGTTAGAAAATAACAAATAGGGTTC
 CGCGCACATTCCCCGAAAAGTGCACCTGACGTCTAAGAACCATATTATCATGACATTAACCTATAAA
 AATAGGCGTATCACGAGGCCCTTCGTC

Coding sequence (SEQ ID NO:149)

ATGAAGGCCATCATCGTGTGCTGATGGTGGTCACAAGCAACGCCATAGAAATCTGTACCGGCATCACCAG
 CAGCAATAGCCCTCACGCGTGAACAGCTACACAGGGCGAAGTGAATGTGACCGCGTGTACCGGC
 CCACAAACACTCAAAGGCCACTTCCGAATCTGAAGGGCACAGAGAACAGGGCAAGCTGTGCCCC
 TGCTGAAATTGACAGATCTGGATGGCTCTGGCAGACCTAAGTGTACAGCAGAACACTCCCTAGCGCCAG
 AGTGTCCATTCTGATGAGTGCACCTGTGACCGAGCGCTGTTTCCATTATGCACGACCCACAATGTGATCA
 TCAGACAGCTGCTTAATCTGCTGAGAGGCTACGAGCACATCAGACTGAGCACCCACAATGTGATCA
 GAAAATGCTCTGGCGGCCCTATAAGATCGGCACATCTGGCAGCTGCCAACATTACAATGGCAATGG
 CTTCTTGCCACCATGGCTGGCGTGCCTAAGAACGATAAGAACAGGCCACCAACCCCTGACAA
 TCGAGGTGCCATATATCTGTACAGAGGGCGAGGATCAGATCACCGTGTGGGATTTCACAGCGACA
 ACACAGATGGCAAGCTGTACGGCGATAGCAAGCCCTCAGAACAGTTTACAGCTGCAATGGCGTAC
 ACACATATGTGCTCAGATCGGCGCTCCCTAATCAGACAGAACATGGCGACTGCCCTAGCTGG
 CGTGGTGGATTACATGGTGCAGAAGTCTGGCAAGACGGGACCATCACATATCAGAGAGGAAT
 CCCAGAGAAGTCTGGCGCTTCTGGAAAGATCAGAACAGCTGATCAAGGGCAGCTGCCCTG
 CGATTGTCGACAGAGAAATACGGCGGCCGAACAAGAGAACGCTTACTACAGGGCGACAGGCCAAG
 CCATCGGCATTGCTTATTGGCTTAAGACCCCTCTGAAGCTGCCAATGGCACAAAGTATAGAACCT
 GCCAAGCTGCTGAAAGAGAGAGGGTTTTGGAGCTATGCCGGCTTTCTGGAAAGGGCGATGGGAGGA
 GATTGCTGGATGGCATGGCTACACATCTCATGGCGCACATGGCGTGGCAGTGGCTGCTGATCT
 CACAGGAAGCCATCAACAAAGATCACCAAGAACCTGACAGCCTGAGCGAGCTGAAAGAAC
 AGACTGTCGGGCCATGGACGACTGCACAAATGAGATCCTGGAACTGGGAGAGAGGCGAC
 AGCGGACATCACAGCAGCCAGATTGAACTGGCTGCTGCTGCTAACGAGGGCATCAT
 ACGAACATCTGCTGGCCCTGGAAAGAACAGCTGAAAGAACATGCTGGGACCTAGGCC
 CGGAGTGGCTGAGACAAAGCACAAGTCAACAGACACCTGCCCTGGATAGAAATT
 CGCGAGTGGCTGAGACATCACATCCGGAGGGCACATCATCAAGCTGCTGA

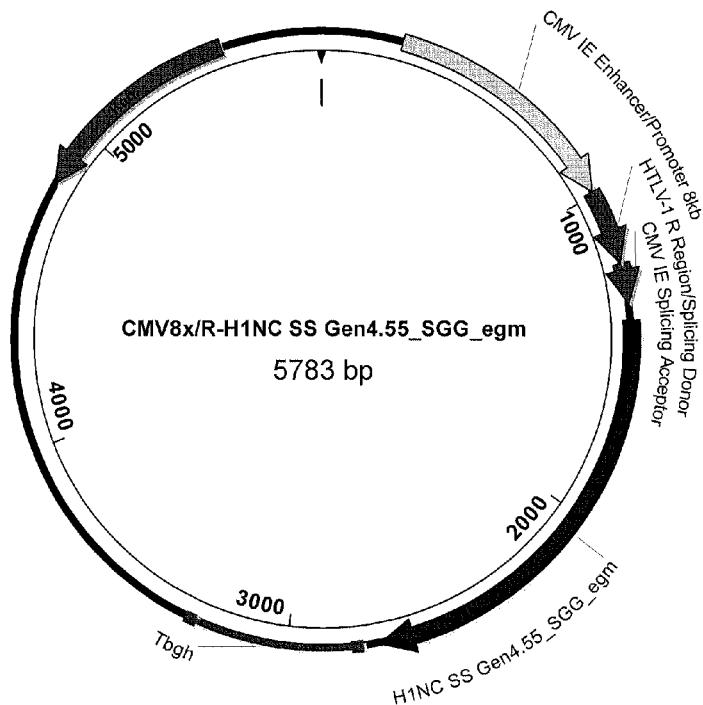
Fig. 34-3

ACGAGCAGGTGAACAAGGAGATGCAGAGCAGCAACCTGTACATGAGCATGAGCAGCTGGTGCTACACCCAC
AGCCTGGACGGCGCCGGCTGTTCTGTCGACCACGCCGAGGAGTACGAGCACGCCAAGAACGCTGAT
CATCTCCTGAACGAGAACAAACGTCGCCCCGTGAGCTGACCAGCATCAGGCCCGAGCACAAAGTCGAGG
GCCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCAGCACATCAGCGAGAGCATCAACAACATCGTG
GACCACGCCATCAAGAGCAAGGACCACGCCACCTCAACTTCCTGCACTGAGCTGATCGGCAAGGAGAACACGGCCTGTACC
GGAGGAGGTGCTGTTCAAGGACATCCTGACAAGATCGAGCTGATCGGCAAGGAGAACACGGCCTGTACC
TGGCCGACCAGTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGGATCC

Translation (SEQ ID NO:68)

MKAIIVLLMVVTNSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPPLTTPTKSHFANLKGTETRGKLCPK
CLNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINA
ENAPGGPYKIGTSGSCPNTNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNE
TQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFNPQTEDGGLPQSGRIVVDYMQKSGKTGTITYQRGILL
PQKVVCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANCKYRPP
AKLLKERGFFFIAIGFLEGGWEGMIAAGWHGTYSHGAHVAVAADLKSTQEAINKITKNLNSLSELEVKNLQ
RLSGAMDELHNEILELDEKVDDLRADETISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGN
GCFETKHKCNCQTCLDRIAAGTFDAGEFSLPTFDSSLNITSGGDIIKLLNEQVNKEMQSSNLYMSMSSWCYTH
SLDGAGLFLFDHAAEYEHAKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIV
DHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLAQYVKGIAKSRSKGS

Fig. 34-4



H1 NC Stabilized Stem Gen4.55_SGG_egm (H1 1999NC SS Gen4.55-ferritin)

Plasmid DNA sequence (SEQ ID NO:150)

TCGCGCGTTGGTGTGATGACGGTGAAACCTCTGACACATGCAGCTCCGGAGACGGTCACAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCGTCAAGGGCGCTCAGGGGGTGTGGCGGGTGTGGGGCTGGCT
TAACTATGCCGCATCAGAGCAGATTGACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCG
TAAGGAGAAAATACCGCATCAGATTGGCTATTGGCATTGCACTACGGTGTATCATAATCATAATATGTACA
TTTATATTGGCTCATGTCCAACATTACGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCA
ATTACGGGAACCTCCATAGCCCATAATGGAGTTCCCGTACATAACTTACGGAATTCCAAACCTGGC
TGACCGCCCAACGACCCCCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAA
CTTCCATTGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTTGGGAATTTCAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGGAATTCCATAAGCTTGCATTATGCCAGTACATGACC
TTATGGGAATTCCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGGTTTG

Fig. 35-1

CGACTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGAACCTCCAAGTCCTCCACCCCCATTGACGCTCA
ATGGGAGTTGTTGACTCACAAAATCAACGGGAATTCCAAAATGTCGTAAACAACCTCGGCCCATG
CGCAAATGGCGGTAGCGTGTACGGTGGGAGGTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATC
GCCTGGAGACGCCATCCACGCTTGTGACCTCATAGAAGACACCAGGACCGATCCAGCCTCATCGCT
CGCATCTCCTTCACGCCGCCCTACCTGAGGCCATCCACGCCGGTTGAGTCGCTCTGCC
GCCTCCCGCTGTGGCTCTGAACCTCGTCCGCCGTAGGTAAGTTAAAGCTCAGGTGAGACCGG
GCCTTGTCGGCGCTCCCTGGAGCCTACCTAGACTCAGCCGGCTCTCACGCTTGTGCTGACCTGCTT
GCTCAACTCTAGTTAACGGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCACCAG
ACATAATAGCTGACAGACTAACAGACTGTTCTCCATGGCTTTCTGAGTCACCGTCGTCGACAGC
TGTGATCAGATATCGCGGCCGCTTAGAGATATCGCACCAGTGAAGGCCAAGCTGCTGGTGTGCTCTGCC
CCTTTACCGCACCTACGCGACACCCTACGCTTGGCTACCGCAACACAGCACCGGACACCCGGTGGAT
ACCGTGCTGGAAAAGAACCGTGAACCGTGAACCCACAGCGTGAACCTGGGATCCGGACTGAGAATGTCACCGG
CCTGAGAACATCCCCAGCATCCAGAGCAGGGCTGTTGGAGGCCATTGCCGGTTTATTGAGGGGGGAT
GGACCGGAATGGTGGATGGGTGGTACCGCTACCAACACCAGAACATGAGCAGGGCTCTGGCTATGCCCGGAT
CAGAAGTCATCCCAGAACGCCATCAACGGCATCACAACAAAGTGAACACCGTGAACAGGAAATGGGCGG
CGATCTGAATGGGACAGAGAGATCAACAACTACACCGCATCATCTACAGCCTGATCAGGAAAGCCAGA
ACCAGCAGAAAACGGCACAGGGCGGATCTGAATGTCAGCAGCAGAACACCTGCTGAGAGCATT
GAGGCCAGCAGCATCTGTCAGCTGACAGTGTGGGCATCAAGCAGCTGAGCAGACCTACAAITGCCAGCT
GCTGGTCTCTGGAAAACGAGAGAACCCCTGGACTTCCACAGCAGAACCGTGAAGAACCTGTACAGGAAAG
TGAAGTCCCAGTGAAGAACACGCCAAGAGATCGCAACGGCTGCTTCGAGTTCTACCAAGTGAAC
AACAGTGCATGAAAGCGTGAAGAACGGCACCTACGACTACCCAAAGTACAGCAGGAAAGCAAGCTGA
CAGAGAGAACATCGACTCGGAGGGCAGCATCAAGCTGCAACGAGCAGGTGAACAAGGAGATGAGA
GCAGCAACCTGTACATGAGCATGAGCAGCTGGTGTACACCCACAGCCTGGACGGCGCCGGCTGTTCTG
TTGACCAAGCCGCGGAGGAGTACAGACGCCAAGAGCTGATCATCTCTGAACCGAGAACACCGTCC
CGTGCAGCTGACCAGCATCAGGCCCGAGACAAGTTGGAGGGCCTGACCCAGATCTCCAGAAGGCC
ACGAGCACGAGCACACATCAGCAGAGCATCAACAAACATCGTGGACCGCCATCAAGAGCAAGGAC
GCCACCTCACTTCTCGCAGTGGTACGTGGCCGAGCAGCACGAGGAGGAGGTGCTGTTCAAGGACATCCT
GGACAAGATCGACTGATCGCAACGAGAACACGGCCTGTACCTGGCCGACAGTACGTGAAGGGCATCG
CCAAGAGCAGAACAGCGGATCTAGCATCATCATATTAGTCTGAAAGGGCGAATTGATCCAGATCT
GCTGTGCTTCTAGTTGCAAGCCATCTGTTGCCCCCTCCCCGTGCTTCTGACCCCTGGAAGGTG
CACTCCCAGTGTCTTCTTAATAAAATGAGGAATTGACATCGCATGTTCTGAGTAGGTGTCATTCTATT
TGGGGGGTGGGGTGGGCAGGACAGCAAGGGGGAGGAATTGAGAACATAGCAGGATGCTGGGATGCG
GTGGGCTCTATGGTACCCAGGTGCTGAAGAATTGACCCGGTTCTCTGGGCCAGAAAGAACAGG
TCCCTCTCTGTGACACACCTGTCAGGCCCTGGTTCTTAGTTCCAGCCCCACTCATAGGACACTCAT
AGCTCAGGAGGGCTCCGCTTCAATCCCACCCCTAAAGTACTTGGAGCGGTCTCTCCCTCCCTCATCAGC
CCACCAAAACCAACCTAGCTCCAAAGAGTGGGAAGAAATTAAAGCAAGATAGGCTATTAGTGCAGAGGG
GAGAAAATGCCCTAACATGTGAGGAAGTAATGAGGAAATCATAGAATTGAGGCATGATTAAAGGCC
ATCATGGGCTTAATCTCCGCTCTCGTCACTGACTCGTCCGCTCGTGTGCTGCGGCCAGCGG
TATCAGCTCACTCAAAGGGCGTAATACGGTTATCCACAGAAATCAGGGATAACGCAGGAAAGAACATGTA
GCAAAAGGCCAGAAAAGGCCAGGAACGGTAAAAAGGCCGGCTTGCTGGCTTTTCCATAGGCTCCGCC
CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGACTATAAGATACC
AGGGCTTCTCCCTGGAAGCTCCCTGTCGGCTCTCTGTCAGGCCCTGCGCTTACGGGATACCTGTC
GCCCTCTCCCTGGAAGCTCCCTGTCGGCTCTCTGTCAGGCCCTGCGCTTACGGGATACCTGTC
CGTTCGCTTCAAGCTGGCTGTGACGAACCCCCCGTTAGCCGACCGTCCGCTTACGGGATACCTGTC
ATCGTCTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGGACTGTAACAGGATTAGC
AGAGCAGGGTATGTAAGGGCTGCTACAGAGTTCTGAAAGTGGGGCTAATACGGCTACACTAGAAGAAC
AGTATTGGTATCTGCGCTCTGTAAGGCCAGTTACCTTGGGAAAGAGTGGTAGCTTGTGATCCGGCA
AACAAACCCAGCTGGTAGCGTGGTTTTGTTGAGCAGCAGGATTAACGGCAGAAAGGGATCT
CAAGAAGATCTTGTATCTACGGGGTCTGACGCTCAGTGGAAAGAAAATCACGTTAAGGGATT
GGTCATGAGATTATCAAAAGGATCTCACCTAGATCTTTAAATTAAAGTGAAGTTAAATCAATCT
AAAGTATATATGAGTAAACCTGGCTGACAGTACCAATGCTTAAATCAGTGAGGCACCTATCTACG
TGTCTATTGCTCATCCATAGTTGCTGACTCGGGGGGGGGGGCGCTAGGGTCTGCTCGTGAAGAAGG
TGTGCTGACTCATACCGCCCTGAATCGCCCCATCATCCAGCAGAAAGTGAAGGGAGGCCACGGTGTG
GAGCTTGTGAGGTGGACAGTGGTAGTTGAACCTTGTGCTTGGCACCGAACGGTCTGCGTTGTG

Fig. 35-2

GGAAGATGCGTGATCTGATCCTCAACTCAGCAAAACTTCGATTATTCAACAAAGCCGGTCCGTCAA
 GTCAGCGTAATGCTCTGCCAGTGTACAACCAATTAAACCATTCTGATTAGAAAAACTCATCGACATCAA
 ATGAAAACCTGCAATTATTATTCATATCAGGATTATCAATACCATTCTGTTGAAAGCGTTCTGTAATGAAG
 GAGAAAACCTCACCGAGGCAGTCCATAGGATGGCAAGATCCTGGTATCGTCTGCGATTCCGACTCGTCCA
 ACATCAATAACACCTATTATTTCCCTCGTCAAAAATAAGGTATCAAGTGAAGAACATACCATGAGTGAC
 GACTGAATCCGGTGAGAATGCCAAAAGCTTATGCATTCTTCAGACTTGTCAACAGGCCAGCATTC
 GCTCGTCATCAAAACTCGCATCAACCAACCGTTATTCAATTGCGATTGCGCTGAGCGAGACGAAT
 ACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGGCAGGAACACTGCCAGCG
 ATCAACAATATTTCACCTGAATCAGGATATTCTCTAATACCTGGAATGCTGTTTCCCGGGATCGCAG
 TGGTAGTAACCATCATCAGGACTACGGATAAAATGCTTGATGGTCAAAGAGGCTATAAATTCCGTC
 AGCCAGTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTGCATGTTGAGAACAA
 CTCTGGCGCATCGGGCTTCCCATAACATCGATAGATTGTCGACCTGATTGCCACATTATCGCAGGCC
 ATTATACCCATAAAATCAGCATCCATGTTGGAAATTAAATCGGGCCTCGAGCAAGACCTTCCCGTTGA
 ATATGGCTCATAAACACCCCTGATTACTGTTTATGTAAGCAGACAGTTTATTGTTCATGATGATATATT
 TTATCTTGTGCAATGTAACATCAGAGATTGAGACACAACGTGGCTTCCCCCCCCCATTATTGAA
 GCATTATCAGGGTTATTGTCATGAGCGGACATATTTGAATGTTAGAAAATAACAAATAGGG
 GTTCCCGCGCACATTCCCCAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
 TAAAAATAGGCGTATCACGGAGGCCCTCGTC

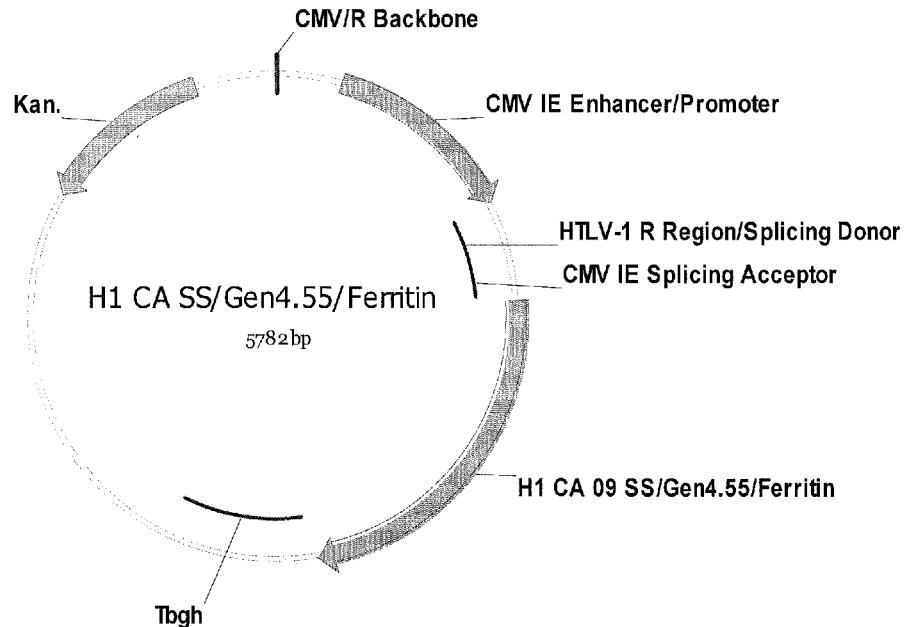
Coding sequence (SEQ ID NO:151)

ATGAAGGCCAAGCTGCTGGTCTGTCACCTTACGCCACCTACGCCACACCATTCGATTGGCTA
 CCACGCCAACACAGCACCACCGTGATACCGCTGGAAAGAACGTCACCGTGACCCACAGCGTGA
 ACTGGGATCCGACTGAGAATGGTCACCGGCTGAGAAACATCCCAGCATCCAGAGCAGAGGCCGTT
 GGAGCCATTGCCGCTTATTGAGGGCGATGGACCGGAATGGGGATGGGGTACGGCTACCACCA
 GAATGAGCAGGGCTCTGGTATGCCCGATCAGAAGTCTACCAAGAACGCATCAACGGCATACCAACA
 AAGTGAACAGCGTGATCGAGAAGATGGCGCGATCCTGAATGGGACAGAGAGATCAACAACACCCAGC
 ATCATCTACAGCCTGATCGAGGAAAGCCAGAACCCAGCAGGAAACGGCACAGGCCGGGATCTGGAATTGT
 GCAGCAGCAGAACACCTGCTGAGGAGCATTGAGGCCAGCAGCATCTGTCAGCTGACAGTGTGGGCA
 TCAAGCAGCTGAGACCTAACATGCCAGCTGCTGGCTCTGGAAAACGAGAGAACCTGGACTTCCAC
 GACAGCAACGTGAAGAACCTGAGGAAAGTGAAGTCCCAGCTGAAGAACACGCCAAGAGATCGGCA
 CGGCTGCTGAGGTTCTACCAAGTCAACACAGACTGAGGAAAGCTGAGGAGCAGCACCTACCGACT
 ACCCAAGTACAGCGAGGAAGCAAGTGAACAGAGAGAAGATCGACTCCGGAGGGCAGATCATCAAGCTG
 CTGAACGAGCAGGTGAACAGAGCAGCACCTGATGAGCAGTGGTCTACAC
 CCACAGCCTGGAGGGCGGCCCTGTTCTGACCAACGCCAGCTGAGGAGTACGGCACGCCAAGAAGC
 TGATCATCTTCTGAAACGAGAACACCTGCCCCGTGAGCTGACAGCATCAGGCCCGAGCACAGTTC
 GAGGGCCTGACCCAGATCTTCCAGAAGGCCAGCAGCACGCCAGCACATCAGCAGAGAGCATCAACACAT
 CGTGGACCAGCCATCAAGAGCAAGGACAGCCACCTCAACTTCTGCACTGAGCTGATGGCAGTGGCGAGCAGC
 ACAGGAGGAGGAGGTGCTGTTCAAGGACATCTGGAAAGATCGAGCTGATGGCAGAAGGAAACCACGGCTG
 TACCTGGCCGACCACTACGTGAAGGGCATGCCAAGAGCAGGAAGAGCGGGATCC

Translation (SEQ ID NO:101)

MKAKLLVLLCTFTATYADTICIGYHANNSTDVTLEKNVTVTHSVNLGGLRMVTGLRNIPSIQSRLF
 GAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYADQKSTQNAINGITNKVNNSIEKMGDPEDREINNYTS
 IIYSLIEESQNQQENGTTGGSGIVQQQNLLRAIEAQHQHLLQLTVWGIKQLQTYNAELLVLLENERTLDFH
 DSNVKNLYEKVKSQQLKNNAKEIGNGCFEFYHKCNMBCMESVKNGTYDYPKYSEESKLNREKIDSGGDIK
 LNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEYEAKKLIIFLNENNVPVOLTSISAPEHKF
 EGLTOIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGL
 YLADQYVKGIAKSRKSGS

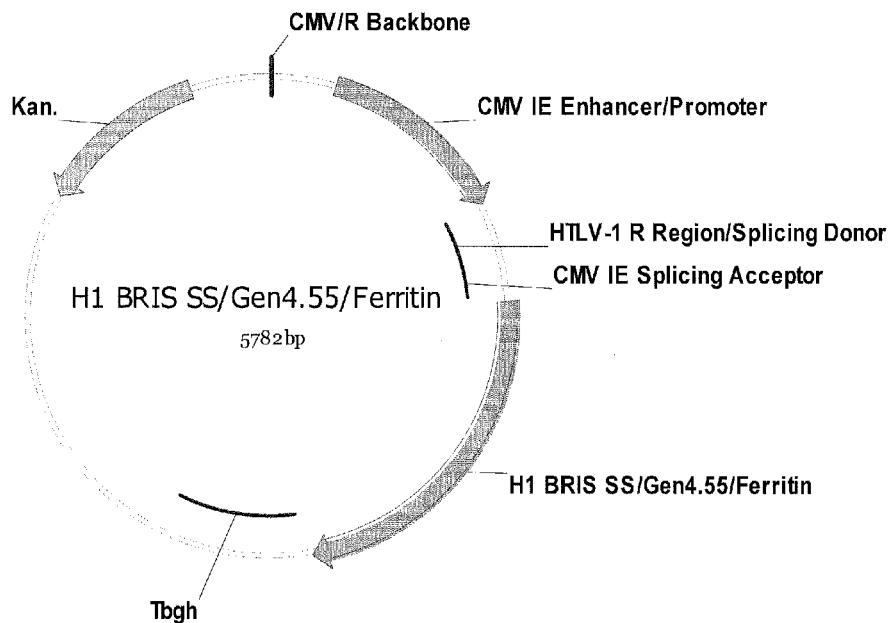
Fig. 35-3



H1 CA SS/Gen4.55/ferritin (SEQ ID NO:152)

Fig. 36-1

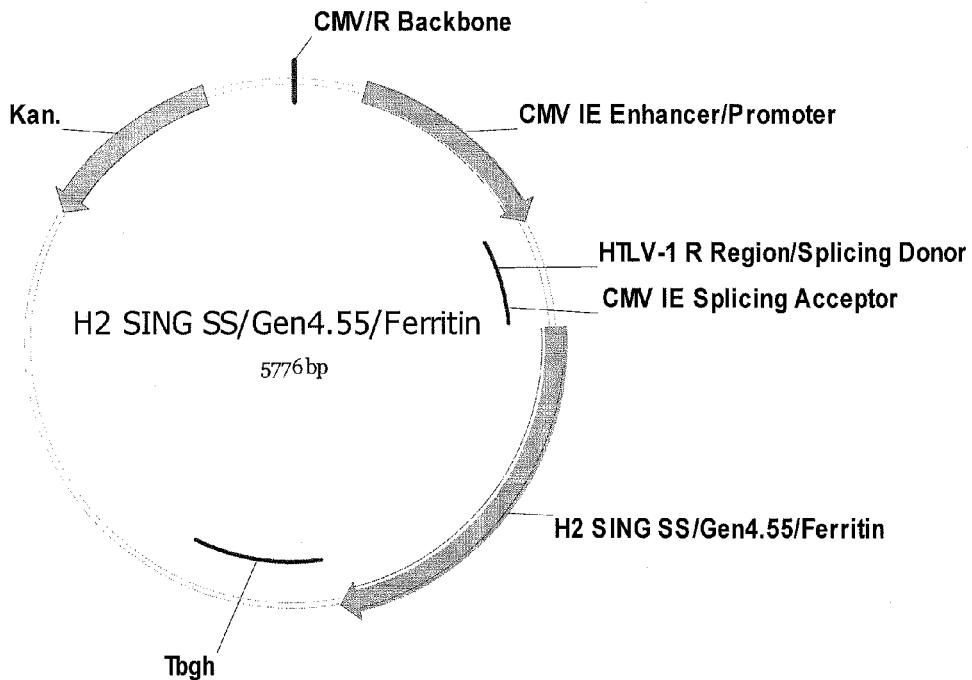
Fig. 36-2



H1 Bris SS/Gen4.55/ferritin (SEQ ID NO:153)

Fig. 37-1

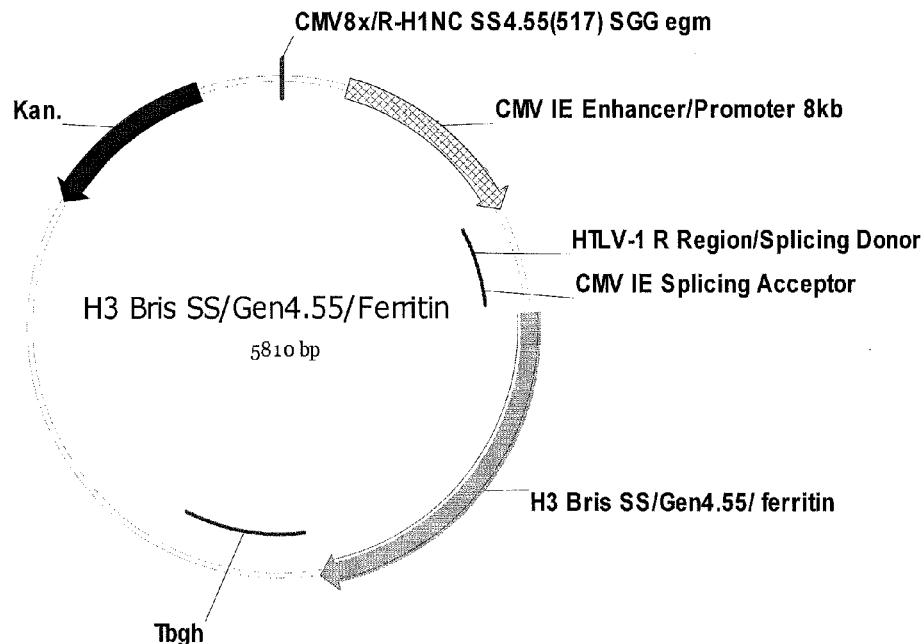
Fig. 37-2



H2 Sing SS/Gen4.55/ferritin (SEQ ID NO:154)

Fig. 38-1

Fig. 38-2

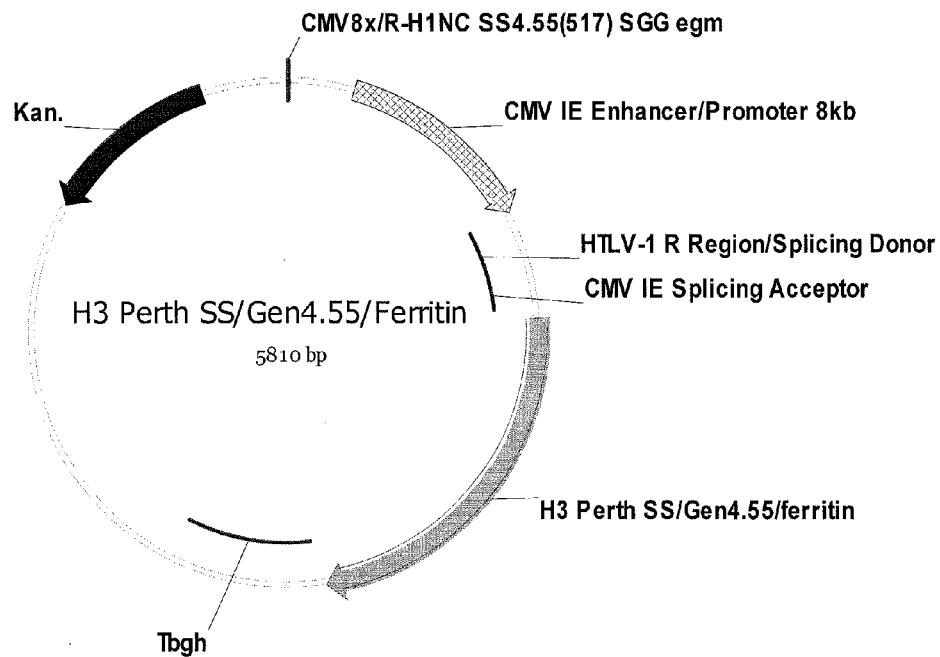


H3 Bris SS/Gen4.55/ferritin (SEQ ID NO:155)

Fig. 39-1

agcatcagcgccccgagcacaagttcagggcctgacccagatcttcagaaggcctacgagcacgcacatcagcgagagcatcaa
caacatcggtgaccacgcccatacgagaagggaccacggccacccataacttcctgcagtggatcgtggcggagcagcacgaggagggtgt
gttcaaggacatctggacaagatcgatcgcaacgagaaccacggcgtacctggcggaccagtagtgcgatcgtgggcatgcctaagag
caggaagagcggatcttagcatcatcatcatcattagtcgaagggcgaattgtatccatgcgtgcctctagttccagccatctgttgtgc
ccctccccgtgccttcgtgaccttggaaagggtgcactccactgtccttcataataatgaggaaattgcacatgcgtgcctgttgtgc
attctattctgggggtgggtggggcaggacagcaaggggggaggattggaaagacaatgcggcatgcgtgggatgcgggtggctctatgg
gtacccagggtgtcaagaattgacccggctcctcggcagaaagaagcaggacatccctctgtgcacacccctgtccacgcccctgg
ttcttagtccagccccactcataggacactcatgcgtcaggaggctccgttcaatccaccggctaaatgtacttggagcgggtctccctcc
tcatcagccaccaaccacccatgcgtcaagatggaaagaaaattaaagcaagataggctttaatgtcagagggagagaaaatgcctc
caacatgtgaggagaatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
cgctgcgtcggtcggtcggtatcgactcaaggcgtaatacggttatccacagaatcagggataacgcaggaa
gaacatgtgaggaaaaggccagcaaaaaggccaggaaaccgtaaaaaggccgtgtcgggtttccataggcgtccgcggccgtac
tcacaaaatgcgtcaagtcagtgagggtggcggaaaccgcacaggactataaagataccaggcgttccctggagctccctgtgcgtct
cctgtccgaccctgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
ggtaggtcgtcgtccaaatgcgtggctgtgcacgaaccccccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
cccgtaagacgacttatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
tggtaggtcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
tccggcaacaaaccaccgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
ctttctacgggtctgcgtcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
attaaaaatgaagtttaatcaatctaataatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
atctgtctattcgatccatcatccatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
gaatcgccccatcatccgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
ccacggaaacgggtctgtcggaaatgtcgatctgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
cagcgtatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
attatcaatccatattttggaaagccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
ctgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
aatccgggtgagaatggcaaaatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
aaccgttatttcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
caggaacactggcagcgcatacaaatatttcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
taaccatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
atcattggcaacgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
attatcgagccatattacccatataatcagcatccatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
acacccctgtattactgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
acgtggcttccccccccccattattgaagcaitattcagggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
taggggtccgcacatttccggaaatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
aggcccttcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc

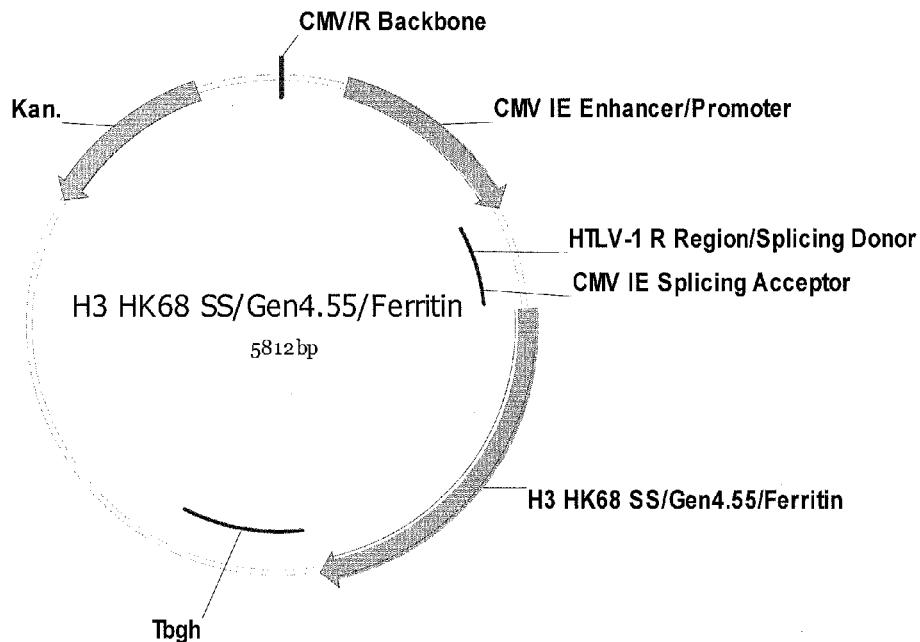
Fig. 39-2



H3 Perth SS/Gen4.55/ferritin (SEQ ID NO:156)

Fig. 40-1

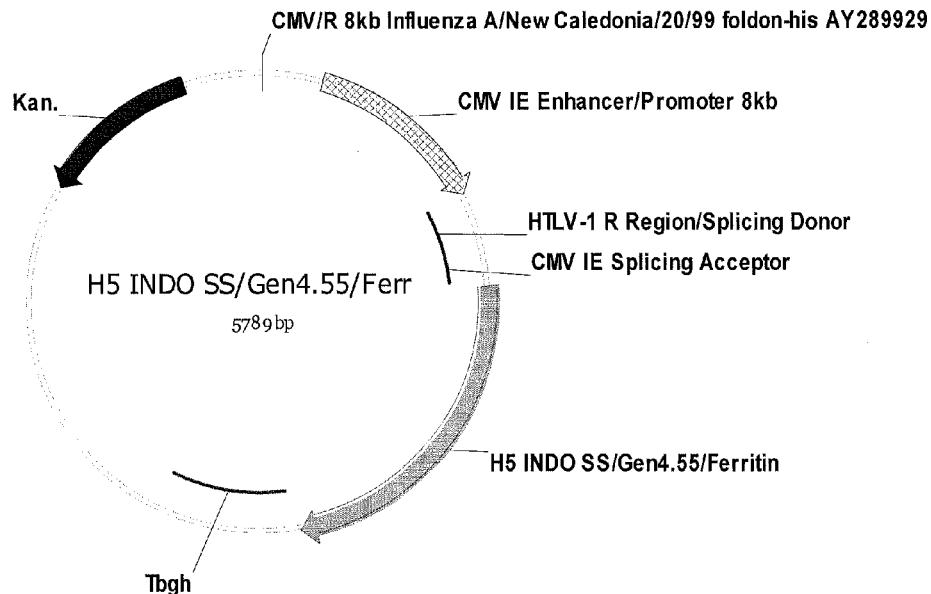
Fig. 40-2



H3 HK SS/Gen4.55/ferritin (SEQ ID NO:157)

Fig. 41-1

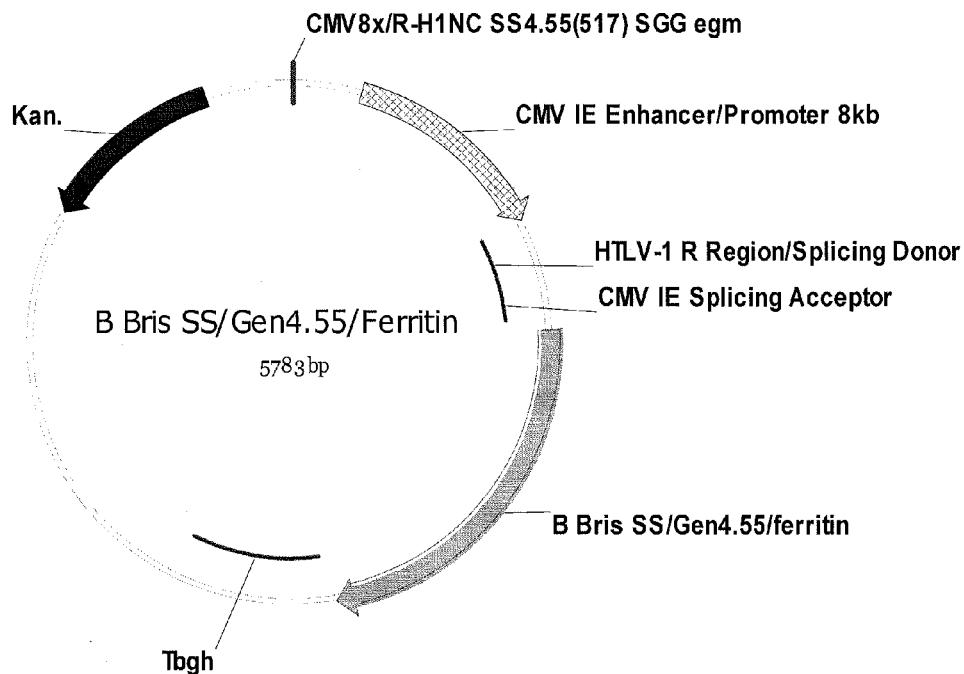
Fig. 41-2



H5 Indo SS/Gen4.55/ferritin (SEQ ID NO:158)

Fig. 42-1

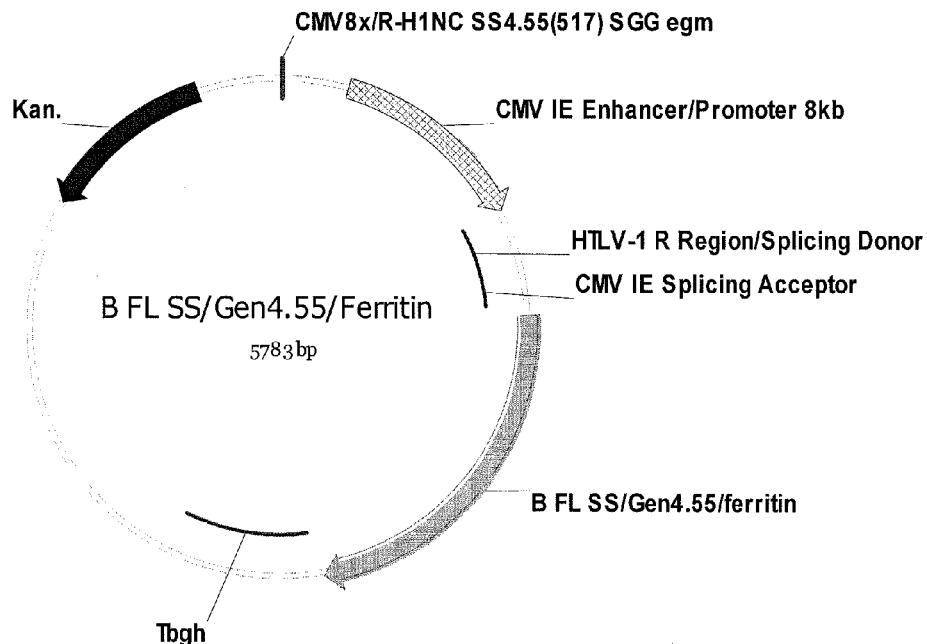
Fig. 42-2



B Bris SS/Gen4.55/ferritin (SEQ ID NO:159)

Fig. 43-1

Fig. 43-2



B FL SS/Gen4.55/ferritin (SEQ ID NO:160)

Fig. 44-1

acccagatcttccagaaggcctacgagcacgcacatcagcgagagcatcaacaacatcgiggaccacgccatcaagagcaaggacca
cgccacctcaacttcgtcagtggtaacgtggccgagcagcagcaggaggagggtgtcaaggacatcctggacaagatcgacgtatcgcc
aacgagaaccacggctgtaccctggccgaccagtaacgtgaaggcatcgccaagagcaggaaagagcgatccatgtttgcctcccccgtgc
cattagtgtaagggcgaattgtccatcgatcgatgttcgttagtgcattcttgccttgccttgccttgccttgccttgccttgc
cactccactgtccttcataaaaatgaggaaattgtcatcgatgttcgttagtgcattcttgccttgccttgccttgccttgc
caagggggaggattggaaagacaatagcaggcatgtgggtatcggtgggtctatggtacccagggtgtgaagaattgaccggcttc
ctggggccagaagaaggcaggcacatccccctctgtgacacacccctgtccacgccccctgtcttagtccaccccactcataggacactcata
gctcaggaggaggctccctcaatcccccctaataacttggagcggtctcccttcatacgccccaccaaaccatcgctccaa
gagtggaaagaaattaaaggcaagataggctattaatgtcagaggagagaaaaatgcctccaacatgtgaggaagtaatgagagaaatcatag
aattttaggccatgatggccatcatggcttaatctccgttcctcgctactgactcgctgcgtcggtcggtcgccgagcggt
atcagctcactcaaaggccgttatccacagaatcagggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca
ggaaccgtaaaaaggccgcgttgccgtttccataggcccccgtacgagcatcacaaaaatgcacgtcaagtgcagggatggcg
aaaccgcacaggactataaagataccaggcgttccccctggaaagctccctgtgcgtctctgttccgaccctgcccattaccggatactgtc
cgcccttcctccctcggaagcggtccctcatcgactcagctgttagtctcgatcggtcggttaggtcggtcgccaaagctggctgtgc
acgaaccccccgttcagccgaccgtcgcccttatccgttaactatcgctgttagtccaaacccgttaagacacgactatcgccactggcagca
gccactggtaacaggattagcagagcgaggatgttagcggtctacagagttctgaagtggcttcaactacggctacactagaagaacag
tatttgtatcgctcgtgaaggcgttaccccgaaaaagagttgttagtcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
tttttgttgcacgcacgatccatcgccgaaaaaaaggatctcaagaagatcttgcgttgcgttgcgttgcgttgcgttgcgttgc
aactcagcttaagggttttgtcatgattatcaaaaaggatcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
tatatgagtaaaacttggctgacagttaccaatgcctaatcgtgaggccatctcagcgtatctgttgcgttgcgttgcgttgc
ggggggggggcgtgaggctcgtaagaaagggtttgtactcataccaggccgtatcgcccatcatccagccagaagtgagg
gagccacgggtgatggagacttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
tgatctgtccctcaactcagccaaatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
aaccaattctgattaaaaactcatcgagcatcaaatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
ctgtatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
cctattaattccctcgtaaaaataaggatctcaatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
ttcttccagactgttcaacaggccagccattacgcgtcgtatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
agacgaaatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
tcacctgaatcaggatattctctaataccctggatgtctgtttccgggatcgatcgatcgatcgatcgatcgatcgatcg
tgcttgatggcgtggaaaggcataaattccgtcagcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
gaaacaactctggcgcacggccatccatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
cagcatccatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
catttacagggttattgtctcatcgagcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
gccacctgacgtctaagaaccattatcatgacattaaacctataaaaataggcgatcgatcgagcccttcgtc

Fig. 44-2

1

**INFLUENZA HEMAGGLUTININ
PROTEIN-BASED VACCINES****CROSS REFERENCE TO RELATED
APPLICATIONS**

This application is a national stage application under 35 U.S.C. 371 and claims the benefit of PCT Application No. PCT/US2012/056822 having an international filing date of Sep. 24, 2012, which designated the United States, which PCT application claimed the benefit of U.S. Provisional Application No. 61/538,663 filed Sep. 23, 2011, and U.S. Provisional Application No. 61/661,209 filed Jun. 18, 2012, the disclosure of each of which are incorporated herein by reference.

REFERENCE TO SEQUENCE LISTING

This application contains a Sequence Listing submitted as an electronic text file named “6137NIAID-26-C1-PCT_sequence_listing_ST25.txt”, having a size in bytes of 338 KB, and created on Sep. 21, 2012. The information contained in this electronic file is hereby incorporated by reference in its entirety pursuant to 37 CFR §1.52(e)(5).

SUMMARY OF THE INVENTION

The present invention provides novel hemagglutinin protein-based influenza vaccines that are easily manufactured, potent, and which elicit broadly neutralizing influenza antibodies. In particular, the present invention provides influenza hemagglutinin proteins, and portions thereof, that are useful in inducing the production of neutralizing antibodies. It also provides novel HA-ferritin nanoparticle (np) vaccines. Such nanoparticles comprise fusion proteins, each of which comprises a monomeric subunit of ferritin joined to an immunogenic portion of an influenza hemagglutinin protein. Because such nanoparticles display influenza hemagglutinin protein on their surface, they can be used to vaccinate an individual against influenza virus.

In one embodiment, the invention is a nanoparticle that comprises a fusion protein, and in this embodiment the fusion protein comprises at least 25 contiguous amino acids from a monomeric ferritin subunit protein joined to a first influenza hemagglutinin (HA) protein, such that the nanoparticle comprises influenza virus HA protein trimers on its surface. The nanoparticle can form an octahedron, which can consist of 24 subunits having 432 symmetry. Further, the monomeric ferritin subunit protein can be selected from a bacterial ferritin, a plant ferritin, an algal ferritin, an insect ferritin, a fungal ferritin and a mammalian ferritin, and in a preferred embodiment, is a *Helicobacter pylori* ferritin protein.

In this embodiment, the monomeric ferritin subunit protein can comprise at least 25 contiguous amino acids of an amino acid sequence selected from SEQ ID NO:2 and SEQ ID NO:5 or can comprise an amino acid at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical to those sequences or can comprise those sequences. In another embodiment, the monomeric subunit comprises a region corresponding to amino acids 5-167 of SEQ ID NO:2.

In this embodiment, the hemagglutinin protein can comprise at least 25 contiguous amino acids from the hemagglutinin protein of an influenza virus selected from A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009

2

(2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), B/Brisbane/60/2008 (2008 Bris, B). Also, the hemagglutinin protein can comprise an amino acid sequence that is selected from the amino acid sequences of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38 or one that is at least 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical thereto. Alternatively, the hemagglutinin protein can comprise an amino acid sequence that is selected from the amino acid sequences of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98 or one that is at least 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical thereto.

In this embodiment, the hemagglutinin protein can be capable of eliciting an immune response to a protein comprising an amino acid sequence selected from SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38 or it can comprise a region selected from a region capable of allowing formation of a hemagglutinin trimer, a stem region, an ectodomain, and a region comprising the amino acid sequence from the amino acid residue immediately distal to the last amino acid of the second helical coiled coil to the amino acid residue proximal to the first amino acid of the transmembrane domain.

The hemagglutinin protein can also comprise a hemagglutinin spike domain, a region corresponding to amino acids 1-519 of SEQ ID NO:8 or an amino acid sequence selected from the group consisting of amino acids 1-519 of SEQ ID NO:8 and SEQ ID NO:11.

In this embodiment, the fusion protein can comprise a linker sequence.

In this embodiment, the nanoparticle can elicit an immune response against a stem region of influenza hemagglutinin, a spike of influenza hemagglutinin, an influenza virus strain that is heterologous to the strain influenza virus from which the hemagglutinin protein was obtained or an influenza virus that is antigenically divergent from the influenza virus from which the hemagglutinin protein was obtained.

In this embodiment, the fusion protein can comprise an amino acid sequence at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical to a sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68, wherein the nanoparticle elicits an immune response against an influenza virus or can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. The fusion protein can also comprise an amino acid sequence at least 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least

about 97% identical, at least about 99% identical to a sequence selected from the group consisting of SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128, wherein the nanoparticle elicits an immune response against an influenza virus.

In this embodiment, the nanoparticle can comprise a second fusion protein comprising a second influenza hemagglutinin protein, wherein the first and second influenza hemagglutinin proteins are from different Types, from different sub-types or different strains of influenza viruses.

Another embodiment of the present invention is a vaccine composition comprising any of the foregoing nanoparticle. The vaccine composition can further comprise at least one additional nanoparticle that comprises at least one hemagglutinin protein from a different strain of influenza than the first hemagglutinin protein and the second hemagglutinin protein.

A further embodiment of the invention is a method to produce a vaccine against influenza virus. The method includes expressing a fusion protein comprising a monomeric ferritin protein joined to an influenza hemagglutinin protein under conditions such that the fusion proteins form a nanoparticle displaying hemagglutinin trimers on its surface and recovering the nanoparticle.

The invention also includes a method to vaccinate an individual against influenza that includes administering a nanoparticle to an individual such that the nanoparticle elicits an immune response against influenza virus. In this embodiment, the nanoparticle comprises a monomeric sub-unit of ferritin joined to an influenza hemagglutinin protein and the nanoparticle displays influenza hemagglutinin trimers on its surface. In this embodiment, the nanoparticle can elicit an immune response to an influenza virus strain that is heterologous to the sub-type or strain of or that is antigenically divergent from the influenza virus from which the hemagglutinin protein was obtained.

This method can further include administering to the individual a first vaccine composition and then at a later time, administering a second vaccine composition comprising a nanoparticle that comprises an HA-SS-ferritin fusion protein. The HA SS-ferritin fusion protein can comprise an amino acid sequence selected from SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98 or one that is at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 97% identical or at least 99% identical thereto, wherein the HA SS-ferritin fusion protein elicits an immune response to an influenza virus. The HA SS-ferritin fusion protein can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128, or one at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 97% identical or at least 99% identical thereto, wherein the HA SS-ferritin fusion protein elicits an immune response to an influenza virus.

In this method, the first vaccine composition can comprise a nanoparticle comprising an ectodomain from the hemagglutinin protein of an influenza virus selected from the group consisting of A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3),

A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), B/Brisbane/60/2008 (2008 Bris, B). Alternatively, the hemagglutinin of the first vaccine composition protein can comprise an amino acid sequence selected from SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38 or one that is at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical thereto. Further, the first vaccine composition can comprise an HA-ferritin fusion protein comprising an amino acid sequence selected from SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68 or an amino acid sequence that is at least 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% identical thereto, wherein the nanoparticle elicits an immune response against an influenza virus.

Administration of the boosting composition is generally weeks or months after administration of the priming composition.

A further embodiment of the present invention is a fusion protein comprising a monomeric ferritin subunit protein joined to an influenza hemagglutinin protein. The monomeric ferritin subunit protein can be selected from a bacterial ferritin, a plant ferritin, an algal ferritin, an insect ferritin, a fungal ferritin and a mammalian ferritin or can be a monomeric subunit of a *Helicobacter pylori* ferritin protein. The monomeric ferritin subunit protein can comprise a domain that allows the fusion protein to self-assemble into nanoparticles. In this embodiment, the monomeric ferritin subunit protein can comprise SEQ ID NO:2 or SEQ ID NO:5 or comprise at least 25 contiguous amino acids from or be at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% to a sequence selected from SEQ ID NO:2 and SEQ ID NO:5 and the fusion protein can be capable of self-assembling into nanoparticles. Additionally, the monomeric subunit can comprise a region corresponding to amino acids 5-167 of SEQ ID NO:2.

In this embodiment, the hemagglutinin protein can comprise at least 25 amino acids from an influenza virus selected from A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), and B/Brisbane/60/2008 (2008 Bris, B). Alternatively, the hemagglutinin protein can be capable of eliciting an immune response to a protein comprising an amino acid sequence selected from SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38 or one that is at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% thereto. In this embodiment, the fusion protein can comprise an amino acid sequence selected from SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59,

5

SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68 or one that is at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% thereto.

Further in this embodiment, the hemagglutinin protein can comprise a region selected from a region capable of allowing trimerization of the hemagglutinin protein, a stem region, an ectodomain, and a region comprising the amino acid sequence from the amino acid residue immediately distal to the last amino acid of the second helical coiled coil to the amino acid residue proximal to the first amino acid of the transmembrane domain. The hemagglutinin protein alternatively can comprise a region corresponding to amino acids 1-519 of SEQ ID NO:8, an amino acid sequence selected from the group consisting of amino acids 1-519 of SEQ ID NO:8 and SEQ ID NO:11, or a hemagglutinin spike domain. Further, the hemagglutinin protein can comprise the stem region from an influenza virus selected from A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), or κ/Brisbane/60/2008 (2008 Bris, B). The hemagglutinin protein can also comprise an amino acid sequence at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% to SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

In this embodiment, the fusion protein can comprise one or more linker sequences or an amino acid sequence of selected from the group consisting of SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128 or a sequence that is at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% thereto.

A further embodiment of the present invention is a nucleic acid molecule encoding any of the fusion proteins described above. In this embodiment, the nucleic acid molecule can be functionally linked to a promoter. Other embodiments of the invention include recombinant cells and viruses that comprise such nucleic acid molecules.

Another embodiment of the invention is a protein comprising an amino acid sequence at least 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 97% identical, at least about 99% to an amino acid selected from SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98, wherein the protein is joined to one or more trimerization domains. In this embodiment, the protein can be joined to at least a portion of the head region of an influenza hemagglutinin protein, comprise one or more linker regions or elicit an immune response against an influenza virus. A further embodiment is a nucleic acid molecule encoding such a protein.

BACKGROUND

Protective immune responses induced by vaccination against influenza virus are primarily directed to the viral

6

hemagglutinin (HA) protein, which is a glycoprotein on the surface of the virus responsible for interaction of the virus with host cell receptors. HA proteins on the virus surface are trimers of hemagglutinin protein monomers that are enzymatically cleaved to yield amino-terminal HA1 and carboxy-terminal HA2 polypeptides. The globular head consists exclusively of the major portion of the HA1 polypeptide, whereas the stem that anchors the hemagglutinin protein into the viral lipid envelope is comprised of 10 HA2 and part of HA1. The globular head of a hemagglutinin protein includes two domains: the receptor binding domain (RBD), an ~148-amino acid residue domain that includes the sialic acid-binding site, and the vestigial esterase domain, a smaller ~75-amino acid residue region just below the RBD. 15 The top part of the RBD adjacent to the 2,6-sialic acid recognition sites includes a large region (amino acids 131-143, 170-182, 205-215 and 257-262, 1918 numbering) (referred to herein as the RBD-A region) of over 6000 Å² per trimer that is 95% conserved between A/South Carolina/1/1918 (1918 SC) and A/California/04/2009 (2009 CA) pandemic strains. The globular head includes several antigenic sites that include immunodominant epitopes. Examples include the Sa, Sb, Ca₁, Ca₂ and Cb antigenic sites (see, for example, Caton A J et al, 1982, *Cell* 31, 417-427). The 20 RBD-A region includes the Sa antigenic site and part of the Sb antigenic site.

Antibodies against influenza often target variable antigenic sites in the globular head of HA, which surround a conserved sialic acid binding site, and thus, neutralize only 30 antigenically closely related viruses. The variability of the HA head is due to the constant antigenic drift of influenza viruses and is responsible for seasonal endemics of influenza. In contrast, gene segments of the viral genome can undergo reassortment (antigenic shift) in host species, creating new viruses with altered antigenicity that are capable of becoming pandemics [Salomon, R. et al. *Cell* 136, 402-410 (2009)]. Until now, each year, influenza vaccine is updated to reflect the predicted HA and neuraminidase (NA) for upcoming circulating viruses.

40 Current vaccine strategies for influenza use either a chemically inactivated or a live attenuated influenza virus. Both vaccines are generally produced in embryonated eggs which present major manufacturing limitations due to the time consuming process and limited production capacity.

45 Another more critical limitation of current vaccines is its highly strain-specific efficacy. These challenges became glaring obvious during emergence of the 2009 H1N1 pandemic, thus validating the necessity for new vaccine platforms capable of overcoming these limitations. Virus-like

50 particles represent one of such alternative approaches and are currently being evaluated in clinical trials [Roldao, A. et al. *Expert Rev Vaccines* 9, 1149-1176 (2010); Sheridan, C. *Nat Biotechnol* 27, 489-491 (2009)]. Instead of embryonated eggs, VLPs that often comprise HA, NA and matrix protein

55 1 (M1) can be mass-produced in mammalian or insect cell expression systems [Haynes, J. R. *Expert Rev Vaccines* 8, 435-445 (2009)]. The advantages of this approach are its particulate, multivalent nature and the authentic display of properly folded, trimeric HA spikes that faithfully mimic the infectious virion. In contrast, by the nature of its assembly, the enveloped VLPs contain a small but finite host cell component that may present potential safety, immunogenicity challenges following repeated use of this platform [Wu, C. Y. et al. *PLoS One* 5, e9784 (2010)]. Moreover, the

60 65 immunity induced by the VLPs is essentially the same as current vaccines do, and thus, does not likely improve both potency and breadth of vaccine-induced protective immu-

nity. In addition to VLPs, a recombinant HA protein has also been evaluated in humans [Treanor, J. J. et al. *Vaccine* 19, 1732-1737 (2001); Treanor, J. J. *JAMA* 297, 1577-1582 (2007)], though the ability to induce protective neutralizing antibody titers are limited. The recombinant HA proteins used in those trials were produced in insect cells and might not form native trimer preferentially [Stevens, J. *Science* 303, 1866-1870 (2004)].

Recently, entirely new classes of broadly neutralizing antibodies against influenza viruses were isolated. One class of antibodies recognizes the highly conserved HA stem [Corti, D. et al. *J Clin Invest* 120, 1663-1673 (2010); Ekiert, D. C. et al. *Science* 324, 246-251 (2009); Kashyap, A. K. et al. *Proc Natl Acad Sci USA* 105, 5986-5991 (2008); Okuno, Y. et al. *J Virol* 67, 2552-2558 (1993); Sui, J. et al. *Nat Struct Mol Biol* 16, 265-273 (2009); Ekiert, D. C. et al. *Science* 333, 843-850 (2011); Corti, D. et al. *Science* 333, 850-856 (2011)], and another class of antibodies precisely recognizes the sialic acid binding site of the RBD on the variable HA head [Whittle, J. R. et al. *Proc Natl Acad Sci USA* 108, 14216-14221 (2011); Krause, J. C. et al. *J Virol* 85, 10905-10908 (2011)]. Unlike strain-specific antibodies, those antibodies are capable of neutralizing multiple antigenically distinct viruses, and hence inducing such antibodies has been a focus of next generation universal vaccine [Nabel, G. J. et al. *Nat Med* 16, 1389-1391 (2010)]. However, robustly eliciting these antibodies with such heterologous neutralizing profile by vaccination has been difficult [Steel, J. et al. *MBio* 1, e0018 (2010); Wang, T. T. et al. *PLoS Pathog* 6, e1000796 (2010); Wei, C. J. et al. *Science* 329, 1060-1064 (2010)].

Despite several alternatives to conventional influenza vaccines, advances in biotechnology in past decades have allowed engineering of biological materials to be exploited for the generation of novel vaccine platforms. Ferritin, an iron storage protein found in almost all living organisms, is an example which has been extensively studied and engineered for a number of potential biochemical/biomedical purposes [Iwahori, K. U.S. Patent 2009/0233377 (2009); Meldrum, F. C. et al. *Science* 257, 522-523 (1992); Naitou, M. et al. U.S. Patent 2011/0038025 (2011); Yamashita, I. *Biochim Biophys Acta* 1800, 846-857 (2010)], including a potential vaccine platform for displaying exogenous epitope peptides [Carter, D. C. et al. U.S. Patent 2006/0251679 (2006); Li, C. Q. et al. *Industrial Biotechnol* 2, 143-147 (2006)]. Its use as a vaccine platform is particularly interesting because of its self-assembly and multivalent presentation of antigen which induces stronger B cell responses than monovalent form as well as induce T-cell independent antibody responses [Bachmann, M. F. et al. *Annu Rev Immunol* 15, 235-270 (1997); Dintzis, H. M. et al. *Proc Natl Acad Sci USA* 73, 3671-3675 (1976)]. Further, the molecular architecture of ferritin, which consists of 24 subunits assembling into an octahedral cage with 432 symmetry has the potential to display multimeric antigens on its surface.

There remains a need for an efficacious influenza vaccine that provides robust protection against influenza virus. There particularly remains a need for an influenza vaccine that protects individuals from heterologous strains of influenza virus, including evolving seasonal and pandemic influenza virus strains of the future. The present invention meets this need by providing a novel HA-ferritin nanoparticle (HA-ferritin np) influenza vaccine that is easily manufactured, potent, and elicits broadly neutralizing influenza antibodies

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1. Molecular design and construction of ferritin particles displaying influenza virus hemagglutinin. (a) Rib-

bon diagram of a subunit of *H. pylori* nonheme ferritin (PDB: 3bve) (left). Amino- and carboxyl-termini are labeled as N and C, respectively. Three ferritin subunits surrounding an octahedral three-fold axis are shown as a ribbon diagram (middle). Residue Asp5 is indicated. The octahedral assembly of the ferritin particle (viewed at 8 Å resolution along an octahedral three-fold axis) and A/Solomon Islands/3/2006 (H1N1) HA trimer (PDB: 3sm5) (viewed down from membrane proximal side) (right). The measured average distance between the Asp5 residues in each ferritin subunit surrounding an octahedral three-fold axis is shown as a triangle. The same equilateral triangle ($a=b=c=28 \text{ \AA}$) is also drawn on the HA trimer (right). (b) Schematic representation of the HA-ferritin expression vector used for protein production. (c) Chromatogram of the size exclusion chromatography of ferritin nanoparticles (np) and HA-np (left). Molecular weights (kDa) of calibration standards are indicated above the curves with vertical lines. Calculated molecular weights for ferritin nanoparticles and HA-np were 419 and 2,165 kDa, respectively, and were within 10% of the predicted molecular weights (408 and 2,040 kDa, respectively). Particle size distribution (radius) of purified ferritin nanoparticles and HA-np was determined by dynamic light scattering (middle). Measured mean diameters (d) are indicated. The polydispersity indices of purified ferritin np and HA-np were 0.035 and 0.011, respectively. Purified HA trimer (thrombin uncleaved), HA-np and ferritin nanoparticles were analyzed by SDS-PAGE (right). (d) Negatively stained transmission electron microscopy images of ferritin nanoparticles (left) and HA-np (right). Images were originally recorded at 67,000 \times magnification. (e) Models representing octahedral four-, three- and two-fold symmetries of HA-ferritin np (top panels) and actual TEM image (bottom panels) are shown. Visible HA spikes are numbered in the images.

FIG. 2. Genetic and structural comparison of ferritins. (a) Phylogenetic tree analysis of ferritins found in RSCB PDB. Twenty-two sequences contain 16 ferritins including Vc (*Vibrio cholerae*), Ec (*E. coli*), Hp (*H. pylori*), Af (*Archaeoglobus fulgidus*), Pf (*Pyrococcus furiosus*), Tm (*Thermatoga maritime*), Pm (*Pseudo-nitzschia multiseries*), Tn (L) (*Trichoplusia ni* light chain), Soybean (chloroplastic), Horse (L) (light chain), Human (L), (H) and (M) (light, heavy chains and mitochondrial, respectively), Mouse (L) (light chain), and Frog (M) and (L) (middle and lower subunits, respectively), and 6 bacterioferritins (B) including Mt (B) (*Mycobacterium tuberculosis*), Pa (B) (*Pseudomonas aeruginosa*), Rs (B) (*Rhodabacter sphaeroides*), Bm (B) (*Brucella melitensis*), Av (B) (*Azobacter vinelandii*), and Ec (B). Protein sequences were aligned using Clustal W2 (www.ebi.ac.uk/Tools/msa/clustalw2) with Gonnet matrix and a phylogenetic tree was generated with the Phylogenodon program (<http://iubio.bio.indiana.edu/treeapp/treeprint-form.html>) using the neighbor-joining method. (b) Comparison of surface exposed residues between *H. pylori* and mouse (light chain) (left) or human (light chain) (middle), and mouse and human (light chains) (right). Conservation of surface exposed residues was rendered by UCSF Chimera using a protein sequence alignment generated by Clustal W2. Conserved and varied residues between the two ferritins are shown as light and dark residues, respectively. PDB files 3bve (*H. pylori*) (left and middle) and 1h96 (mouse light chain) (right) were used for surface rendering.

FIG. 3. Antigenic characterization of HA-ferritin np. (a) Binding of mAbs directed to globular head and stem of HA was measured by ELISA. Equal amount (200 ng of HA per

well) of HA trimer (\blacktriangle), TIV (\blacksquare), HA-ferritin (\bullet) or Ferr (equimolar amount as HA-Ferr) (\circ) were coated on the plates and wells were probed with anti-head mAb (3u-u) and anti-stem mAb CR6261. The half maximal effective concentrations (EC_{50}) of binding were calculated for each antibody and showed as ng ml^{-1} (b) Inhibition of antibody-mediated neutralization of 1999 NC pseudotyped virus by using HA trimer, HA-Ferr or Ferr as a competitor. Inhibition of neutralization was plotted as percent inhibition respect to no competitor control. The anti-stem neutralizing mAbs, F10 (left) and CR6261 (right) were used at 3.125 and 25 $\mu\text{g ml}^{-1}$, respectively. Competitor proteins were added to the reactions at a final concentration of 20 $\mu\text{g ml}^{-1}$.

FIG. 4. Immune responses in HA-np-immunized mice. (a) HAI (left), IC_{90} neutralization (middle), and anti-HA ab endpoint titers (right) against 1999 NC HA after two immunizations with 0.17 μg (amount of H1 HA) of TIV or HA-np with or without Ribi adjuvant and a 3-week interval. The immune sera were collected 2 weeks after the second immunization. The data are presented as box-and-whiskers plots (boxed from lower to upper quartile with whiskers from minimum to maximum) with lines at the mean ($n=5$). (b) Neutralization breadth of the immune sera elicited by HA-trimer, TIV, or HA-np. An additional group of mice ($n=4$) was immunized twice with 20 μg of trimeric HA protein using Ribi adjuvant and a 4-week interval. The immune sera were collected 2 weeks after the second immunization. IC_{50} neutralization titers against a panel of H1N1 pseudotyped viruses were determined. (c) Cellular and humoral immune responses against *H. pylori* (top) and mouse (bottom) ferritins. Mice were immunized twice with 1.67 μg (amount of H1 HA) of TIV or HA-np, or 0.57 μg of ferritin nanoparticles (equimolar to HA-np) using Ribi adjuvant and a 3-week interval. The splenocytes and immune sera were harvested 11 days after the second immunization. Cytokine-producing CD4 $^+$ and CD8 $^+$ T cells were measured by ICS (left), and ab response was detected by ELISA (right). All cells expressing IFN- γ , TNF α , or IL-2 were identified as cytokine $^+$ cells. The percentage of cytokine $^+$ cells in CD4 $^+$ and CD8 $^+$ T cells that were activated in response to stimulation with specific peptides covering the entire *H. pylori* or mouse ferritins (heavy and light chains combined) were plotted. Recombinant *H. pylori* and purified mouse (liver) ferritins were used for detecting anti-ferritin ab responses. The data are presented as box-and-whiskers plots with lines at the mean ($n=5$).

FIG. 5. Successive immunization of HA-nanoparticles in mice. Mice were pre-immunized with 1.67 μg (amount of HA) of 2009 Perth (H3) HA-nanoparticles or 0.57 μg (equimolar to HA-nanoparticle) of empty ferritin nanoparticles at week 0 and then immunized with 1.67 μg (amount of HA) of 1999 NC (H1) HA-nanoparticles at week 3. Ribi was used as an adjuvant. Another group of mice was immunized with 1999 NC (H1) HA-nanoparticles without pre-immunization of empty ferritin nanoparticles or H3 HA-nanoparticles. (a) Ab responses to *H. pylori* ferritin (left) and 2009 Perth H3 HA (right). Immune sera collected 2 weeks after the immunization with H3 HA-nanoparticles or empty ferritin nanoparticles were analyzed by ELISA. (b) Immune responses to 1999 NC (H1) after 1999 NC (H1) HA-nanoparticle immunization. Naïve mice or mice with pre-immunity to ferritin or H3 HA were immunized with H1 HA-nanoparticles at week 3 and HAI (left), IC_{90} neutralization (middle) and ELISA (right) Ab titers were measured 2 weeks after the immunization. The data are presented as box-and-whiskers plots with lines at the mean ($n=5$).

FIG. 6. Development of trivalent HA-np. (a) HA-np consisting of HAs from 2009 CA (H1), 2009 Perth (H3) or 2006 FL (type B) were purified and visualized by TEM. (b) HAI titers against 2009 CA (H1N1) and 2009 Perth (H3N2) viruses in immunized mice. Mice were immunized twice with 1.67 μg (amount of HA) of monovalent H1, monovalent H3, monovalent type B, or 5.0 μg (total amount of HA) of trivalent HA-np or TIV (2011-2012 season) using Ribi adjuvant with a 3-week interval. Immune sera were collected 2 weeks after the second immunization. The data are presented as box-and-whiskers plots with lines at the mean ($n=5$).

FIG. 7. Protective immunity induced in ferrets immunized with the HA-np. Ferrets were immunized twice with PBS (control), 7.5 μg (2.5 μg of H1 HA) of TIV or 2.5 μg (amount of HA) 1999 NC HA-np using Ribi adjuvant and a 4-week interval. Control animals received PBS. (a) HAI (left), IC_{90} neutralization (middle), and anti-HA ab endpoint titers (right) in immunized ferrets against homologous 1999 NC HA were determined. Immune sera were collected 3 and 2 weeks after the first (R. Salomon, R. G. Webster, The influenza virus enigma. *Cell* 136, 402-410 (2009) and second (L. C. Lambert, A. S. Fauci, Influenza vaccines for the future. *N Engl J Med* 363, 2036-2044 (2010)) immunizations, respectively. The data are presented as box and whisker plots with lines at the mean ($n=6$). (b) Protection of immunized ferrets from an unmatched 2007 Bris virus challenge. Ferrets were challenged with $10^{6.5}$ 50% egg infectious dose (EID $_{50}$) of 2007 Bris virus 5 weeks after the second immunization. Virus titers in the nasal washes from 1, 3 and 5 days post challenge were determined by a 50% tissue culture infectious dose (TCID $_{50}$) assay (left). One of six ferrets in the TIV-immunized group showed measurable virus on day 5. Virus titers in 4 out of 6 ferrets on day 3 and 6 out of 6 ferrets on day 5 in the HA-np-immunized group were under the detection limit (<102). The mean viral loads with standard deviation (s.d.) at each time point were plotted ($n=6$). Change in the body weight after the virus challenge was also monitored (right). Each data point represents the mean percent change in body weight from the pre-challenge (day 0). The mean body weight changes with standard error (s.e.) at each time point were plotted ($n=6$).

FIG. 8. Improved neutralization breadth and detection of stem- and RBS-directed abs. (a) Neutralization breadth of immune sera in ferrets. IC_{50} neutralization titers against a panel of H1N1 pseudotyped viruses (left) and HAI titers against 1934 PR8 and 2007 Bris H1N1 viruses (right) were determined. The HAI titers are presented as box-and-whiskers plots with lines at the mean ($n=6$). (b) Stem- and RBS-directed abs elicited by HA-np immunization. Ferret immune sera (diluted 1:100) were pre-absorbed with Δ Stem HA-expressing cells and their binding to WT or Δ Stem HA were analyzed by ELISA (left). The immune sera (diluted 1:1,000) were pre-absorbed with Δ RBS HA-expressing cells and their binding to WT or Δ RBS HA were analyzed by ELISA (middle). The mean endpoint dilution titers were plotted with s.d. ($n=6$). Competition ELISA with stem-directed mAb CR6261 (right). The immune sera pre-absorbed with Δ Stem were tested for binding to HA in the presence of an isotype control IgG or CR6261. Each symbol represents the titer of an individual ferret ($n=6$). (c) Neutralization competition with WT, Δ Stem or Δ RBS HA protein (left). The neutralization of HA-np immune sera against 1986 Sing, 1995 Beijing, 1999 NC and 2007 Bris was measured in the presence of irrelevant protein (control), WT, Δ Stem or Δ RBS HA as a competitor. Percent neutralizations at serum dilution 1:200 (1986 Sing, 2007 Bris), 1:800 (1995

11

Beijing) or 1:3,200 (1999 NC) were plotted. Each symbol represents the individual ferret serum and mean is indicated as a red line with s.d. (n=6 for 1986 Sing, 1995 Beijing and 1999 NC; n=3 for 2007 Bris). The relative contribution of the stem- and RBS-directed neutralization was determined by the inhibition of neutralization for each competitor protein (right). Mean percent contributions in neutralizing each virus were plotted as pile-up bars (n=6).

FIG. 9. Characterization of ΔRBS HA probe. (a) Crystal structure of HA (A/Solomon Islands/3/2006) complex with an anti-RBS mAb CH65 Fab (PDB: 3sm5) (J. R. Whittle et al., Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc Natl Acad Sci USA* 108, 14216-14221 (2011)) (left). Close up view of CH65 contact area (right). The residue HA1 190 which has been mutated to be glycosylated in ΔRBS mutant is highlighted. The CH65 Fab-bound HA1 protomer is darkened. (b) Characterization of the soluble trimer of WT and ΔRBS HAs from 1999 NC and 2007 Bris. The WT and ΔRBS HA proteins were immunoprecipitated with anti-RBS (CH65), stem (CR6261) and control (anti-HIV, VRC01) mAbs. Immune complexes were then dissolved in Laminin buffer and analyzed by SDS-PAGE. Antibody heavy and light chains are labeled as HC and LC, respectively.

FIG. 10. Purification of HA-np. HA-np were purified by routine iodixanol gradient ultracentrifugation routinely. Fractions containing HA-np were confirmed by SDS-PAGE and Western blotting using a mAb against 1999 NC HA. The HA-np were enriched in the fraction with density of ~1.15 g/ml.

FIG. 11. Protocol for immunization of mice and ferrets using pan-group 1 HA-ferritin np. Mice were injected intramuscularly twice (Week 0 and week 4) with PBS (control) or 6.8 ug (1.7 ug of each HA-ferritin np) pan-group 1 vaccine in Ribi. Ferrets were injected intramuscularly twice (Week 0 and week 4) with PBS (control) or 10 ug (2.5 ug of each HA-ferritin np) pan-group 1 vaccine in Ribi.

FIG. 12. Neutralization activity of mouse antisera against Group 1 HA pseudotyped viruses. Neutralization activity of murine antisera from control or pan Group1 HA-np immunized mice against the indicated HA pseudotyped viruses. IC50 titers are shown for all panels.

FIG. 13. Neutralization activity of ferret antisera against Group1 HA pseudotyped viruses. Neutralization activity of ferret antisera from control or pan Group1 HA np immunized ferrets against the indicated HA pseudotyped viruses. IC50 titers are shown for all panels.

FIG. 14. H1 HAI assays were performed using the sera obtained from the ferritin immunization studies. These studies were performed using actual H1 virus, and H2 and H5 HAI were performed using HA-ferritin np

FIG. 15. Protection of ferrets from viral challenge with Influenza A/Brisbane/59/2007 Brisbane (H1N1) (2007 Bris). Two groups of ferrets (n=6 for control and n=5 for pan-group1 immune) were immunized with pan Group1 HA np vaccine or PBS (control) and challenged with heterologous 2007 Bris virus ($10^{6.5}$ EID₅₀). Virus titers were measured in nasal swabs collected on day 3 and day 5 post challenge. Titters were determined using end-point titration in MDCK cells.

FIG. 16. Protection of ferrets from viral challenge with Influenza A/Mexico/2009 (H1N1) (2009 Mex). Two groups of ferrets (n=6) were immunized with pan Group1 HA np vaccine or PBS (control) and challenged with heterologous 2009 Mex virus ($10^{6.5}$ EID₅₀). Virus titers were measured in

12

nasal swabs collected on day 3 and day 5 post challenge. Titers were determined using end-point titration in MDCK cells.

FIG. 17. Conservation of the influenza HA stem region. (left, right) Neutralizing antibodies that react with both Group 1 and Group 2 viruses act at the sites of vulnerability shown in the Figure. (Right) Space filling model of influenza HA protein illustrating amino acid sequence conservation in over 800 human H1N1 strains. Light residues indicate residues that are 100% conserved. Dark residues as indicate sites of variation.

FIG. 18. Design of HA Stabilized Stem protein. (A) Schematic of the HA SS (bottom) in comparison to HA (top). HA SS was constructed by inserting a GWG linker between residues 42 and 314 of HA1 RBD head, a gp41 post-fusion trimerization motif inserted in place of residues 59 through 93 of HA2, a GG linker between HA2 and the gp41 HR2 helix and an NGTGGGSG linker between the two gp41 helices. The gene sequence of H1 NC 99 SS is provided in the supplemental materials. (B) Trimeric and monomeric representation of HA (PDB entry 1RU7) in comparison to the HA SS model. Coloring is respective to above panel, with the monomeric representation also illustrating the CR6261 epitope as yellow and HA residues which are omitted in the stabilized HA stem as grey. (C) CR6261, FI6v3, and the germline of the VH1-69 Ab 70-5B03 have similar affinity to HA and SS by ELISA. HA SS competes with CR6261 (D) binding to HA and (E) neutralization of H1N1 pseudovirus similar to soluble HA trimer.

FIG. 19. Size exclusion chromatogram of HA and HA SS probes. Calibration standards are shown above the curves as vertical lines.

FIG. 20. Electron microscopic analysis of nanoparticles. Purified SS-np were subjected to transmission electron microscopic analysis. The samples were negatively stained with ammonium molybdate and images were recorded on a Tecnai T12 microscope (FEI) at 80 kV with a CCD camera (AMT Corp.). Images of lower (left) and higher (right) magnifications are shown. The SS spikes were protruding perpendicularly from the particle core and clearly visible.

FIG. 21. Antigenic characterization of HA SS-ferritin np. The ability of purified HA SS and HA SS-np to bind to monoclonal Abs CR6261 (left) and FI6v3 (right) was characterized by ELISA. HA and HIV gp120 proteins served as controls.

FIG. 22. Immune sera of mice immunized heterologously with HA-np prime and HA SS-np boost are reactive to the conserved HA stem epitope. Antibodies elicited by vaccination target the conserved HA stem epitope as individual mice possess differential binding (a minimum of 2-fold difference in endpoint dilution) between wt and Δstem HA variants. The percentage of mice displaying differential binding is given above matched wt and Δstem constructs. Error bars represent standard error.

FIG. 23. Immune sera of mice immunized with HA SS neutralizes diverse pseudovirus stains. IC50 values are shown for individual mice against H1 homosubtypic strains and H2, H5 and H9 group-1 heterosubtypic strains. Dashed lines represents the lowest dilution assayed (50). Error bars represent standard error.

FIG. 24. Boosting with HA SS-np increases neutralizing titers in ferrets against H1N1 New Calendonia. Pseudovirus neutralizing titers were calculated for preimmune, HA FL-np primed, and HA SS-np boosted sera from individual mice. Error bars represent standard deviation of values.

13

FIG. 25. Map and sequence of CMV8x/R-H1NC HA(517)_SGG-egm (SEQ ID NO:130), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:131) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 26. Map and sequence of CMV8x/R-H1CA HA(518)_SGG-egm (SEQ ID NO:132), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:133) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 27. Map and sequence of CMV8x/R-H2Sing HA(514)_SGG-egm (SEQ ID NO:134), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:135) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 28. Map and sequence of CMV8x/R-H3HK HA(519)_SGG-egm (SEQ ID NO:136), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:137) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 29. Map and sequence of CMV8x/R-H3Bris HA(519)_SGG-egm (SEQ ID NO:138), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:139) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 30. Map and sequence of CMV8x/R-H5Indo HA(520)_SGG-egm (SEQ ID NO:140), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:141) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 31. Map and sequence of CMV8x/R-B.Florida HA(534)_SGG-egm (SEQ ID NO:142), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:143) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 32. Map and sequence of CMV8x/R-H3Perth HA(519)_SGG-egm (SEQ ID NO:144), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:145) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 33. Map and sequence of CMV8x/R-H1Bris HA(517)_SGG-egm (SEQ ID NO:146) and the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:147) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 34. Map and sequence of CMV8x/R-B.Bris HA(535)_SG G-egm (SEQ ID NO:148), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:149) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 35. Map and sequence of CMV8x/R-H1NC SS Gen4.55_SGG-egm (SEQ ID NO:150), the nucleic acid sequence encoding the related HA-ferritin fusion protein (SEQ ID NO:151) and the amino acid sequence of the encoded HA-ferritin fusion protein.

FIG. 36. Map and sequence of CMV/R H1 CA SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 37. Map and sequence of CMV/R H1 Bris SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 38. Map and sequence of CMV/R H2 Sing SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 39. Map and sequence of CMV/R H3 Bris SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 40. Map and sequence of CMV/R H3 Perth SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 41. Map and sequence of CMV/R H3 HK68 SS/Gen4.55/ferritin (SEQ ID NO:152).

14

FIG. 42. Map and sequence of CMV/R H5 Indo SS/Gen4.55/ferritin (SEQ ID NO:152).

FIG. 43. Map and sequence of CMV/R B Bris SS/Gen4.55/ferritin (SEQ ID NO:152).

5 FIG. 44. Map and sequence of CMV/R B FL SS/Gen4.55/ferritin (SEQ ID NO:152).

DETAILED DESCRIPTION OF THE INVENTION

10 The present invention relates to a novel vaccine for influenza virus. More specifically, the present invention relates to novel, influenza hemagglutinin protein-based vaccines that elicit an immune response against a broad range 15 of influenza viruses. It also relates to self-assembling ferritin-based, nanoparticles that display immunogenic portions of influenza hemagglutinin protein on their surface. Such nanoparticles are useful for vaccinating individuals against influenza virus. Accordingly, the present invention also 20 relates to fusion proteins for producing such nanoparticles and nucleic acid molecules encoding such proteins. Additionally, the present invention relates to, methods of producing nanoparticles of the present invention, and methods of using such nanoparticles to vaccinate individuals.

25 Before the present invention is further described, it is to be understood that this invention is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and 30 is not intended to be limiting, since the scope of the present invention will be limited only by the claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. For 35 example, a nucleic acid molecule refers to one or more nucleic acid molecules. As such, the terms "a", "an", "one or more" and "at least one" can be used interchangeably. Similarly the terms "comprising", "including" and "having" can be used interchangeably. It is further noted that the 40 claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

45 In addition to the above, unless specifically defined otherwise, the following terms and phrases, which are common to the various embodiments disclosed herein, are defined as follows:

As used herein, the term immunogenic refers to the ability 50 of a specific protein, or a specific region thereof, to elicit an immune response to the specific protein, or to proteins comprising an amino acid sequence having a high degree of identity with the specific protein. According to the present invention, two proteins having a high degree of identity have 55 amino acid sequences at least 80% identical, at least 85% identical, at least 87% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical.

As used herein, an immune response to a vaccine, or 60 nanoparticle, of the present invention is the development in a subject of a humoral and/or a cellular immune response to a hemagglutinin protein present in the vaccine. For purposes of the present invention, a "humoral immune response" refers to an immune response mediated by antibody molecules, including secretory (IgA) or IgG molecules, while a "cellular immune response" is one mediated by T-lymphocytes and/or other white blood cells. One important aspect of

15

cellular immunity involves an antigen-specific response by cytolytic T-cells ("CTL's). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by the major histocompatibility complex (MHC) and expressed on the surfaces of cells. CTLs help induce and promote the destruction of intracellular microbes, or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves an antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function, and focus the activity of, nonspecific effector cells against cells displaying peptide antigens in association with MHC molecules on their surface. A cellular immune response also refers to the production of cytokines, chemokines and other such molecules produced by activated T-cells and/or other white blood cells, including those derived from CD4+ and CD8+ T-cells.

Thus, an immunological response may be one that stimulates CTLs, and/or the production or activation of helper T-cells. The production of chemokines and/or cytokines may also be stimulated. The vaccine may also elicit an antibody-mediated immune response. Hence, an immunological response may include one or more of the following effects: the production of antibodies (e.g., IgA or IgG) by B-cells; and/or the activation of suppressor, cytotoxic, or helper T-cells and/or T-cells directed specifically to a hemagglutinin protein present in the vaccine. These responses may serve to neutralize infectivity, and/or mediate antibody-complement, or antibody dependent cell cytotoxicity (ADCC) to provide protection to an immunized individual. Such responses can be determined using standard immunoassays and neutralization assays, well known in the art.

According to the present invention all nomenclature used to classify influenza virus is that commonly used by those skilled in the art. Thus, a Type, or Group, of influenza virus refers to influenza Type A, influenza Type B or influenza type C. It is understood by those skilled in the art that the designation of a virus as a specific Type relates to sequence difference in the respective M1 (matrix) protein or NP (nucleoprotein). Type A influenza viruses are further divided into Group 1 and Group 2. These Groups are further divided into subtypes, which refers to classification of a virus based on the sequence of its HA protein. Examples of current commonly recognized subtypes are H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 or H16. Group 1 influenza subtypes are H1, H2, H5, H7 and H9. Group 2 influenza subtypes are H4, H4, H6, H8, H10, H11, H12, H13, H14, H15 and H16. Finally, the term strain refers to viruses within a subtype that differ from one another in that they have small, genetic variations in their genome.

As used herein, neutralizing antibodies are antibodies that prevent influenza virus from completing one round of replication. As defined herein, one round of replication refers the life cycle of the virus, starting with attachment of the virus to a host cell and ending with budding of newly formed virus from the host cell. This life cycle includes, but is not limited to, the steps of attaching to a cell, entering a cell, cleavage and rearrangement of the HA protein, fusion of the viral membrane with the endosomal membrane, release of viral ribonucleoproteins into the cytoplasm, formation of new viral particles and budding of viral particles from the host cell membrane.

As used herein, broadly neutralizing antibodies are antibodies that neutralize more than one type, subtype and/or strain of influenza virus. For example, broadly neutralizing antibodies elicited against an HA protein from a Type A influenza virus may neutralize a Type B or Type C virus. As a further example, broadly neutralizing antibodies elicited

16

against an HA protein from Group I influenza virus may neutralize a Group 2 virus. As an additional example, broadly neutralizing antibodies elicited against an HA protein from one sub-type or strain of virus, may neutralize another sub-type or strain of virus. For example, broadly neutralizing antibodies elicited against an HA protein from an H1 influenza virus may neutralize viruses from one or more sub-types selected from the group consisting of H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 or H16.

As used herein, an influenza hemagglutinin protein, or HA protein, refers to a full-length influenza hemagglutinin protein or any portion thereof, that is capable of eliciting an immune response. Preferred HA proteins are those that are capable of forming a trimer. An epitope of a full-length influenza hemagglutinin protein refers to a portion of such protein that can elicit a neutralizing antibody response against the homologous influenza strain, i.e., a strain from which the HA is derived. In some embodiments, such an epitope can also elicit a neutralizing antibody response against a heterologous influenza strain, i.e., a strain having an HA that is not identical to that of the HA of the immunogen.

With regard to hemagglutinin proteins, it is understood by those skilled in the art that hemagglutinin proteins from different influenza viruses may have different lengths due to mutations (insertions, deletions) in the protein. Thus, reference to a corresponding region refers to a region of another protein that is identical, or nearly so (e.g., at least 95%, 30 identical, at least 98% identical or at least 99% identical), in sequence, structure and/or function to the region being compared. For example, with regard to the stem region of a hemagglutinin protein, the corresponding region in another hemagglutinin protein may not have the same residue numbers, but will have a nearly identical sequence and will perform the same function. To better clarify sequences comparisons between viruses, numbering systems are used by those in the field, which relate amino acid positions to a reference sequence. Thus, corresponding amino acid residues in hemagglutinin proteins from different strains of influenza may not have the same residue number with respect to their distance from the N-terminal amino acid of the protein. For example, using the H3 numbering system, reference to residue 100 in A/New Caledonia/20/1999 (1999 NC, H1) does not mean it is the 100th residue from the N-terminal amino acid. Instead, residue 100 of A/New Caledonia/20/1999 (1999 NC, H1) aligns with residue 100 of influenza H3N2 strain. The use of such numbering systems is understood by those skilled in the art. Unless otherwise noted, reference to amino acids in hemagglutinin proteins herein is made using the H3 numbering system.

According to the present invention, a trimerization domain is a series of amino acids that when joined (also referred to as fused) to a protein or peptide, allow the fusion protein to interact with other fusion proteins containing the trimerization domain, such that a trimeric structure is formed. Any known trimerization domain can be used in the present invention. Examples of trimerization domains include, but are not limited to, the HIV-1 gp41 trimerization domain, the SIV gp41 trimerization domain, the Ebola virus gp-2 trimerization domain, the HTLV-1 gp-21 trimerization domain, the T4 fibrin trimerization domain (i.e., foldon), the yeast heat shock transcription factor trimerization domain, and the human collagen trimerization domain.

As used herein, a variant refers to a protein, or nucleic acid molecule, the sequence of which is similar, but not identical to, a reference sequence, wherein the activity of the

variant protein (or the protein encoded by the variant nucleic acid molecule) is not significantly altered. These variations in sequence can be naturally occurring variations or they can be engineered through the use of genetic engineering technique known to those skilled in the art. Examples of such techniques are found in Sambrook J, Fritsch E F, Maniatis T et al., in Molecular Cloning—A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, 1989, pp. 9.31-9.57, or in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6, both of which are incorporated herein by reference in their entirety.

With regard to variants, any type of alteration in the amino acid, or nucleic acid, sequence is permissible so long as the resulting variant protein retains the ability to elicit neutralizing antibodies against an influenza virus. Examples of such variations include, but are not limited to, deletions, insertions, substitutions and combinations thereof. For example, with regard to proteins, it is well understood by those skilled in the art that one or more (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or 10), amino acids can often be removed from the amino and/or carboxy terminal ends of a protein without significantly affecting the activity of that protein. Similarly, one or more (e.g., 2, 3, 4, 5, 6, 7, 8, 9 or 10) amino acids can often be inserted into a protein without significantly affecting the activity of the protein.

As noted, variant proteins of the present invention can contain amino acid substitutions relative to the influenza HA proteins disclosed herein. Any amino acid substitution is permissible so long as the activity of the protein is not significantly affected. In this regard, it is appreciated in the art that amino acids can be classified into groups based on their physical properties. Examples of such groups include, but are not limited to, charged amino acids, uncharged amino acids, polar uncharged amino acids, and hydrophobic amino acids. Preferred variants that contain substitutions are those in which an amino acid is substituted with an amino acid from the same group. Such substitutions are referred to as conservative substitutions.

Naturally occurring residues may be divided into classes based on common side chain properties:

- 1) hydrophobic: Met, Ala, Val, Leu, Ile;
- 2) neutral hydrophilic: Cys, Ser, Thr;
- 3) acidic: Asp, Glu;
- 4) basic: Asn, Gln, His, Lys, Arg;
- 5) residues that influence chain orientation: Gly, Pro; and
- 6) aromatic: Trp, Tyr, Phe.

For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class.

In making amino acid changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. The hydropathic indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5). The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte et al., 1982, *J. Mol. Biol.* 157:105-31). It is known that certain amino acids may be substituted for other amino acids having a similar hydrophatic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic

indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological invention, as in the present case. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, i.e., with a biological property of the protein. The following hydrophilicity values have been assigned to these amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 ± 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity.

Desired amino acid substitutions (whether conservative or non-conservative) can be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the HA protein, or to increase or decrease the immunogenicity, solubility or stability of the HA proteins described herein. Exemplary amino acid substitutions are shown below in Table 1.

TABLE 1

Amino Acid Substitutions	
Original Amino Acid	Exemplary Substitutions
Ala	Val, Leu, Ile
Arg	Lys, Gln, Asn
Asn	Gln
Asp	Glu
Cys	Ser, Ala
Gln	Asn
Glu	Asp
Gly	Pro, Ala
His	Asn, Gln, Lys, Arg
Ile	Leu, Val, Met, Ala
Leu	Ile, Val, Met, Ala
Lys	Arg, Gln, Asn
Met	Leu, Phe, Ile
Phe	Leu, Val, Ile, Ala, Tyr
Pro	Ala
Ser	Thr, Ala, Cys
Thr	Ser
Trp	Tyr, Phe
Tyr	Trp, Phe, Thr, Ser
Val	Ile, Met, Leu, Phe, Ala

As used herein, the phrase significantly affect a protein's activity refers to a decrease in the activity of a protein by at least 10%, at least 20%, at least 30%, at least 40% or at least 50%. With regard to the present invention, such an activity may be measured, for example, as the ability of a protein to elicit neutralizing antibodies against an influenza virus. Such activity may be measured by measuring the titer of such antibodies against influenza virus, or by measuring the number of types, subtypes or strains neutralized by the

19

elicited antibodies. Methods of determining antibody titers and methods of performing virus neutralization assays are known to those skilled in the art. In addition to the activities described above, other activities that may be measured include the ability to agglutinate red blood cells and the binding affinity of the protein for a cell. Methods of measuring such activities are known to those skilled in the art.

As used herein, a fusion protein is a recombinant protein containing amino acid sequence from at least two unrelated proteins that have been joined together, via a peptide bond, to make a single protein. The unrelated amino acid sequences can be joined directly to each other or they can be joined using a linker sequence. As used herein, proteins are unrelated, if their amino acid sequences are not normally found joined together via a peptide bond in their natural environment(s) (e.g., inside a cell). For example, the amino acid sequences of monomeric subunits that make up ferritin, and the amino acid sequences of influenza hemagglutinin proteins are not normally found joined together via a peptide bond.

The terms individual, subject, and patient are well-recognized in the art, and are herein used interchangeably to refer to any human or other animal susceptible to influenza infection. Examples include, but are not limited to, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, seals, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The terms individual, subject, and patient by themselves, do not denote a particular age, sex, race, and the like. Thus, individuals of any age, whether male or female, are intended to be covered by the present disclosure and include, but are not limited to the elderly, adults, children, babies, infants, and toddlers. Likewise, the methods of the present invention can be applied to any race, including, for example, Caucasian (white), African-American (black), Native American, Native Hawaiian, Hispanic, Latino, Asian, and European. An infected subject is a subject that is known to have influenza virus in their body.

As used herein, a vaccinated subject is a subject that has been administered a vaccine that is intended to provide a protective effect against an influenza virus.

As used herein, the terms exposed, exposure, and the like, indicate the subject has come in contact with a person or animal that is known to be infected with an influenza virus.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates, which may need to be independently confirmed.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

20

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. All combinations of the embodiments are specifically embraced by the present invention and are disclosed herein just as if each and every combination was individually and explicitly disclosed. In addition, all sub-combinations are also specifically embraced by the present invention and are disclosed herein just as if each and every such sub-combination was individually and explicitly disclosed herein.

According to the present invention, vaccines are provided that elicit a broad immune response against influenza viruses. Some vaccines disclosed herein may elicit an immune response against the entire HA protein, while others may elicit an immune response against a specific region or portion of an influenza HA protein. Moreover, the inventors have discovered that specific fusion proteins comprising portions of hemagglutinin protein are useful for eliciting a broad immune response against influenza viruses. Each of these embodiments will now be disclosed in detail below.

Vaccines Against the Stem Region of Influenza HA Protein

As stated previously, the amino acid sequence of the stem region of the hemagglutinin protein is highly conserved across types, sub-types and strains of influenza viruses and contains a site of vulnerability for group 1 viruses. Thus, an immune response directed this region of the HA protein may protect individuals against influenza viruses from several types, sub-types and/or strains.

Consequently, one embodiment of the present invention is a protein that elicits an immune response against the stem region of an influenza HA protein. In one embodiment, the immune response can be directed against the stem region of an HA protein from a virus selected from the group consisting of influenza A viruses, influenza B viruses and influenza C viruses. In one embodiment, the immune response can be directed against the stem region of an HA protein from a virus selected from the group consisting of an H1 influenza virus, an H2 influenza virus, an influenza H3 virus, an influenza H4 virus, an influenza H5 virus, an influenza H6 virus, an H7 influenza virus, an H8 influenza virus, an H9 influenza virus, an H10 influenza virus, an H11 influenza virus, an H12 influenza virus, an H13 influenza virus, an H14 influenza virus, an H15 influenza virus and an H16 influenza virus. In one embodiment, the immune response can be directed against the stem region of an HA protein from a strain of virus selected from the group of viruses listed in Table 2.

TABLE 2

SEQ ID NO	Comments	FERRITIN	
1	Coding sequence for ferritin monomeric subunit protein from <i>H. pylori</i>		
2	Amino acid sequence encoded by SEQ ID NO: 1		
3	Complement of SEQ ID NO 1		
4	Nucleic acid sequence encoding amino acids 5-167 from SEQ ID NO: 2; Asn19 has been replaced with Gln		
5	Amino acid sequence encoded by SEQ ID NO: 3		
6	Complement of SEQ ID NO 3		

TABLE 2-continued

SEQ ID NO	Comments
FULL LENGTH HA	
7	Nucleic acid sequence encoding full length hemagglutinin protein from A/New Caledonia/20/1999 (1999 NC, H1)(GenBank: AY289929)
8	Amino acid sequence encoded by SEQ ID NO: 7 (full length hemagglutinin protein from A/New Caledonia/20/1999 (1999 NC, H1)(GenBank: AY289929))
9	Complement of SEQ ID NO: 7
	ECTODOMAINS
10	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/New Caledonia/20/1999 (1999 NC, H1).
11	Amino acid sequence encoded by SEQ ID NO: 10 (ectodomain from hemagglutinin protein from A/New Caledonia/20/1999 (1999 NC, H1). Amino acids 1-517 from SEQ ID NO: 8.
12	Complement of SEQ ID NO: 10
13	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/California/04/2009 (2009 CA, H1).
14	Amino acid sequence encoded by SEQ ID NO: 13 (ectodomain from hemagglutinin protein from A/California/04/2009 (2009 CA, H1))
15	Complement of SEQ ID NO: 13
16	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Singapore/1/1957 (1957 Sing, H2).
17	Amino acid sequence encoded by SEQ ID NO: 16 (ectodomain from hemagglutinin protein from A/Singapore/1/1957 (1957 Sing, H2))
18	Complement of SEQ ID NO: 16
19	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Hong Kong/1/1968 (1968 HK, H3).
20	Amino acid sequence encoded by SEQ ID NO: 19) ectodomain from hemagglutinin protein from A/Hong Kong/1/1968 (1968 HK, H3))
21	Complement of SEQ ID NO: 19
22	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Brisbane/10/2007 (2007 Bris, H3).
23	Amino acid sequence encoded by SEQ ID NO: 22 (ectodomain from hemagglutinin protein from A/Brisbane/10/2007 (2007 Bris, H3))
24	Complement of SEQ ID NO: 22.
25	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Indonesia/05/2005 (2005 Indo, H5)
26	Amino acid sequence encoded by SEQ ID NO: 25 (ectodomain from hemagglutinin protein from A/Indonesia/05/2005 (2005 Indo, H5))
27	Complement of SEQ ID NO: 25
28	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from B/Florida/4/2006 (2006 Flo, B)
29	Amino acid sequence encoded by SEQ ID NO: 28 (ectodomain from hemagglutinin protein from B/Florida/4/2006 (2006 Flo, B))
30	Complement of SEQ ID NO: 28
31	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Perth/16/2009 (2009 Per, H3)
32	Amino acid sequence encoded by SEQ ID NO: 31 (ectodomain from hemagglutinin protein from A/Perth/16/2009 (2009 Per, H3))
33	Complement of SEQ ID NO: 31
34	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from A/Brisbane/59/2007 (2007 Bris, H1)
35	Amino acid sequence encoded by SEQ ID NO: 34 (ectodomain from hemagglutinin protein from A/Brisbane/59/2007 (2007 Bris, H1))
36	Complement of SEQ ID NO: 34
37	Nucleic acid sequence encoding ectodomain from hemagglutinin protein from B/Brisbane/60/2008 (2008 Bris, B)

TABLE 2-continued

SEQ ID NO	Comments
5	Amino acid sequence encoded by SEQ ID NO: 37 (ectodomain from hemagglutinin protein from B/Brisbane/60/2008 (2008 Bris, B))
38	Complement of SEQ ID NO: 37
	FERRITIN-HA ECTODOMAIN FUSION
10	Nucleic acid sequence encoding SEQ ID NO: 41
40	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/New Caledonia/20/1999 (1999 NC, H1))
41	Complement of SEQ ID NO: 40
15	Nucleic acid sequence encoding SEQ ID NO: 44
42	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/California/04/2009 (2009 CA, H1))
43	Complement of SEQ ID NO: 43
44	Nucleic acid sequence encoding SEQ ID NO: 47
45	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/Singapore/1/1957 (1957 Sing, H2))
46	Complement of SEQ ID NO: 46
47	Nucleic acid sequence encoding SEQ ID NO: 50
20	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/Hong Kong/1/1968 (1968 HK, H3))
48	Complement of SEQ ID NO: 49
49	Nucleic acid sequence encoding SEQ ID NO: 53
50	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/Brisbane/10/2007 (2007 Bris, H3))
25	Complement of SEQ ID NO: 51
51	Nucleic acid sequence encoding SEQ ID NO: 53
52	Amino acid sequence of ferritin-HA fusion (ectodomain from hemagglutinin protein from A/Indonesia/05/2005 (2005 Indo, H5))
53	Complement of SEQ ID NO: 52
30	Nucleic acid sequence encoding SEQ ID NO: 56
54	Amino acid sequence of ferritin-HA fusion protein (ectodomain from hemagglutinin protein from B/Florida/4/2006 (2006 Flo, B))
55	Complement of SEQ ID NO: 55
56	Nucleic acid sequence encoding SEQ ID NO: 59
57	Amino acid sequence of ferritin-HA fusion protein (ectodomain from hemagglutinin protein from A/Perth/16/2009 (2009 Per, H3))
58	Complement of SEQ ID NO: 61
59	Nucleic acid sequence encoding SEQ ID NO: 62
35	Amino acid sequence of ferritin-HA fusion protein (ectodomain from hemagglutinin protein from A/Brisbane/59/2007 (2007 Bris, H1))
60	Complement of SEQ ID NO: 64
61	Nucleic acid sequence encoding SEQ ID NO: 68
62	Amino acid sequence of ferritin-HA fusion protein (ectodomain from hemagglutinin protein from A/Perth/16/2009 (2009 Per, H3))
40	Complement of SEQ ID NO: 66
63	Nucleic acid sequence encoding SEQ ID NO: 68
64	Amino acid sequence of ferritin-HA fusion protein (ectodomain from hemagglutinin protein from B/Brisbane/60/2008 (2008 Bris, B))
65	Complement of SEQ ID NO: 67
	STEM REGION
50	Nucleic acid molecule encoding SEQ ID NO: 71
70	Amino acid sequence of stem region from A/New Caledonia/20/1999 (1999 NC, H1)(GenBank: AY289929)
71	Complement of SEQ ID NO: 70
72	Nucleic acid sequence encoding SEQ ID NO: 74
73	Amino acid sequence of stem region from A/California/04/2009 (2009 CA, H1)
74	Complement of SEQ ID NO: 73
75	Nucleic acid sequence encoding SEQ ID NO: 77
76	Amino acid sequence of stem region from A/Singapore/1/1957 (1957 Sing, H2)
77	Complement of SEQ ID NO: 76
60	Nucleic acid sequence encoding SEQ ID NO: 80
78	Amino acid sequence of stem region from A/Hong Kong/1/1968 (1968 HK, H3)
79	Complement of SEQ ID NO: 79
80	Nucleic acid sequence encoding SEQ ID NO: 83
81	Amino acid sequence of stem region from A/Brisbane/10/2007 (2007 Bris, H3)
82	Complement of SEQ ID NO: 83
65	Nucleic acid sequence encoding SEQ ID NO: 83
83	Amino acid sequence of stem region from A/Brisbane/10/2007 (2007 Bris, H3)

TABLE 2-continued

SEQ ID NO	Comments
84	Complement of SEQ ID NO: 82
85	Nucleic acid sequence encoding SEQ ID NO: 86
86	Amino acid sequence of stem region from A/Indonesia/05/2005 (2005 Indo, H5)
87	Complement of SEQ ID NO: 85
88	Nucleic acid sequence encoding SEQ ID NO: 89
89	Amino acid sequence of stem region from B/Florida/4/2006 (2006 Flo, B)
90	Complement of SEQ ID NO: 88
91	Nucleic acid sequence encoding SEQ ID NO: 92
92	Amino acid sequence of stem region from A/Perth/16/2009 (2009 Per, H3)
93	Complement of SEQ ID NO: 91
94	Nucleic acid sequence encoding SEQ ID NO: 95
95	Amino acid sequence of stem region from A/Brisbane/59/2007 (2007 Bris, H1)
96	Complement of SEQ ID NO: 94
97	Nucleic acid sequence encoding SEQ ID NO: 98
98	Amino acid sequence of stem region from B/Brisbane/60/2008 (2008 Bris, B)
99	Complement of SEQ ID NO: 97
FERRITIN- HA STEM REGION FUSION	
100	Nucleic acid sequence encoding SEQ ID NO: 101
101	Amino acid sequence of ferritin-HA stem region fusion protein A/New Caledonia/20/1999 (1999 NC, H1)
102	Complement of SEQ ID NO: 100
103	Nucleic acid sequence encoding SEQ ID NO: 104
104	Amino acid sequence of ferritin-HA stem region fusion protein (H1 CA)
105	Complement of SEQ ID NO: 103
106	Nucleic acid sequence encoding SEQ ID NO: 107
107	Amino acid sequence of ferritin-HA stem region fusion protein (H2 Sing)
108	Complement of SEQ ID NO: 106
109	Nucleic acid sequence encoding SEQ ID NO: 110
110	Amino acid sequence of ferritin-HA stem region fusion protein (H3 Hong Kong)
111	Complement of SEQ ID NO: 109
112	Nucleic acid sequence encoding SEQ ID NO: 113
113	Amino acid sequence of ferritin-HA stem region fusion protein (H5 Indonesia)
114	Complement of SEQ ID NO: 112
115	Nucleic acid sequence encoding SEQ ID NO: 116
116	Amino acid sequence of ferritin-HA stem region fusion protein (A/Brisbane/59/2007 (2007 Bris, H1))
117	Complement of SEQ ID NO: 115
118	Nucleic acid sequence encoding SEQ ID NO: 119
119	Amino acid sequence of ferritin-HA stem region fusion protein (A/Brisbane/10/2007 (2007 Bris, H3))
120	Complement of SEQ ID NO: 118
121	Nucleic acid sequence encoding SEQ ID NO: 122
122	Amino acid sequence of ferritin-HA stem region fusion protein (A/Perth/16/2009 (2009 Per, H3))
123	Complement of SEQ ID NO: 121
124	Nucleic acid sequence encoding SEQ ID NO: 125
125	Amino acid sequence of ferritin-HA stem region fusion protein (B/Brisbane/60/2008 (2008 Bris, B))
126	Complement of SEQ ID NO: 124
127	Nucleic acid sequence encoding SEQ ID NO: 128
128	Amino acid sequence of ferritin-HA stem region fusion protein (B/Florida/4/2006 (2006 Flo, B))
129	Complement of SEQ ID NO: 127
130	Sequence of plasmid CMV8x/R-H1NC HA(517)_SGG_ egn Synthetic sequence (FIG. 25)
131	Nucleic acid sequence encoding SEQ ID NO: 41. Contains stop codon. Identical to SEQ ID NO: 40, Synthetic (FIG. 25)
132	Sequence of plasmid CMV8x/R-H1CA HA(518)_SGG_ egn Synthetic sequence (FIG. 26)
133	Nucleic acid sequence encoding SEQ ID NO: 44. Nearly identical to SEQ ID NO: 43 but lacks stop codon. (FIG. 26)
134	Sequence of plasmid CMV8x/R-H2SINGHA(514)_SGG_ egn, Synthetic sequence (FIG. 27)

TABLE 2-continued

SEQ ID NO	Comments
5	Nucleic acid sequence encoding SEQ ID NO: 47. Nearly identical to SEQ ID NO: 46 but lacks stop codon, Synthetic (FIG. 27)
135	Sequence of plasmid CMV8x/R-H3HK HA(519)_SGG_ egn Synthetic sequence (FIG. 28)
136	Nucleic acid sequence encoding SEQ ID NO: 50. Nearly identical to SEQ ID NO: 49 but lacks stop codon. Synthetic (FIG. 28)
137	Sequence of plasmid CMV8x/R-H3Bris HA(519)_SGG_ egn Synthetic sequence (FIG. 29)
10	Nucleic acid sequence encoding SEQ ID NO: 53. Nearly identical to SEQ ID NO: 52 but lacks stop codon. Synthetic (FIG. 29)
138	Sequence of plasmid CMV8x/R-H3Indo HA(520)_SGG_ egn, Synthetic sequences (FIG. 30)
139	Nucleic acid sequence encoding SEQ ID NO: 56. Nearly identical to SEQ ID NO: 55 but lacks stop codon. Synthetic (FIG. 30)
15	Sequence of plasmid CMV8x/R-H1Florda HA(534)_SGG_ egn, Synthetic sequence (FIG. 31)
140	Nucleic acid sequence encoding SEQ ID NO: 59. Nearly identical to SEQ ID NO: 58 but lacks stop codon. Synthetic (FIG. 31)
141	Sequence of plasmid CMV8x/R-H3-Perth HA(519)_SGG_ egn, Synthetic sequence (FIG. 32)
20	Nucleic acid sequence encoding SEQ ID NO: 62. Nearly identical to SEQ ID NO: 61 but lacks stop codon. Synthetic (FIG. 32)
142	Sequence of plasmid CMV8x/R-H1Bris HA(517)_SGG_ egn Synthetic sequence (FIG. 33)
143	Nucleic acid sequence encoding SEQ ID NO: 65. Nearly identical to SEQ ID NO: 64 but lacks stop codon. Synthetic (FIG. 33)
144	Sequence of plasmid CMV8x/R-H3-Perth HA(519)_SGG_ egn Synthetic sequence (FIG. 32)
25	Nucleic acid sequence encoding SEQ ID NO: 67 but lacks stop codon. Synthetic (FIG. 34)
145	Sequence of plasmid CMV8x/R-H1NC SS Gen4.55_SGG_ egn Synthetic sequence (Fig. 35)
146	Nucleic acid sequence encoding SEQ ID NO: 101. Identical to SEQ ID NO: 100. Both lack stop codon. (FIG. 35)
147	Sequence of plasmid CMV8x/R-H1CA SS/Gen4.55/Ferritin Synthetic sequence (FIG. 36)
30	Sequence of plasmid H1Bris SS/Gen4.55/Ferritin Synthetic sequence (FIG. 37)
148	Sequence of plasmid H1Sing SS/Gen4.55/Ferritin Synthetic sequence (FIG. 38)
149	Sequence of plasmid H3Bris SS/Gen4.55/Ferritin Synthetic sequence (FIG. 39)
35	Sequence of plasmid H3 HK68 SS/Gen4.55/Ferritin Synthetic sequence (FIG. 41)
150	Sequence of plasmid H5Indo SS/Gen4.55/Ferritin Synthetic sequence (FIG. 42)
151	Sequence of plasmid H1Perth SS/Gen4.55/Ferritin Synthetic sequence (FIG. 40)
40	Sequence of plasmid H3 HK68 SS/Gen4.55/Ferritin Synthetic sequence (FIG. 41)
152	Sequence of plasmid H1Bris SS/Gen4.55/Ferritin Synthetic sequence (FIG. 37)
153	Sequence of plasmid H1Sing SS/Gen4.55/Ferritin Synthetic sequence (FIG. 38)
154	Sequence of plasmid H3Bris SS/Gen4.55/Ferritin Synthetic sequence (FIG. 39)
45	Sequence of plasmid H3 HK68 SS/Gen4.55/Ferritin Synthetic sequence (FIG. 42)
155	Sequence of plasmid H1Perth SS/Gen4.55/Ferritin Synthetic sequence (FIG. 40)
156	Sequence of plasmid H3 HK68 SS/Gen4.55/Ferritin Synthetic sequence (FIG. 41)
157	Sequence of plasmid H5Indo SS/Gen4.55/Ferritin Synthetic sequence (FIG. 42)
50	Sequence of plasmid H1Perth SS/Gen4.55/Ferritin Synthetic sequence (FIG. 43)
158	Sequence of plasmid B Bris SS/Gen4.55/Ferritin Synthetic sequence (FIG. 44)
159	Sequence of plasmid B FL SS/Gen4.55/Ferritin Synthetic sequence Synthetic sequence (FIG. 44)
160	Sequence of plasmid B FL SS/Gen4.55/Ferritin Synthetic sequence Synthetic sequence (FIG. 44)

55 One type of immune response is a B-cell response, which results in the production of antibodies against the antigen that elicited the immune response. Thus, one embodiment of the present invention is a protein that elicits antibodies that bind to the stem region of influenza HA protein from a virus selected from the group consisting of influenza A viruses, influenza B viruses and influenza C viruses. One embodiment of the present invention is a protein that elicits antibodies that bind to the stem region of influenza HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an influenza H3 virus HA protein, an influenza H4 virus HA protein, an

25

influenza H5 virus HA protein, an influenza H6 virus HA protein, an H7 influenza virus HA protein, an H8 influenza virus HA protein, an H9 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein and an H16 influenza virus HA protein. One embodiment of the present invention is a protein that elicits antibodies that bind to the stem region of influenza HA protein from a strain of virus selected from the viruses listed in Table 2.

While all antibodies are capable of binding to the antigen which elicited the immune response that resulted in antibody production, preferred antibodies are those that neutralize an influenza virus. Thus, one embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to the stem region of influenza HA protein from a virus selected from the group consisting of influenza A viruses, influenza B viruses and influenza C viruses. One embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to the stem region of influenza HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an influenza H3 virus HA protein, an influenza H4 virus HA protein, an influenza H5 virus HA protein, an influenza H6 virus HA protein, an H7 influenza virus HA protein, an H8 influenza virus HA protein, an H9 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein and an H16 influenza virus HA protein. One embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to the stem region of influenza HA protein from a strain of virus selected from the viruses listed in Table 2. One embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to a protein comprising an amino acid sequence at least 80% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. One embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

Neutralizing antibodies elicited by proteins of the present invention can neutralize viral infections by affecting any step in the life cycle of the virus. For example, neutralizing antibodies may prevent an influenza virus from attaching to a cell, entering a cell, releasing viral ribonucleoproteins into the cytoplasm, forming new viral particles in the infected cell and budding new viral particles from the infected host cell membrane. In one embodiment, neutralizing antibodies elicited by proteins of the present invention prevent influenza virus from attaching to the host cell. In one embodiment, neutralizing antibodies elicited by proteins of the present invention prevent influenza virus from entering the host cell. In one embodiment, neutralizing antibodies elicited by proteins of the present invention prevent fusion of viral membranes with endosomal membranes. In one embodiment, neutralizing antibodies elicited by proteins of the present invention prevent release of ribonucleoproteins into the cytoplasm of the host cell. In one embodiment, neutralizing antibodies elicited by proteins of the present

26

invention prevent assembly of new virus in the infected host cell. In one embodiment, neutralizing antibodies elicited by proteins of the present invention prevent release of newly formed virus from the infected host cell.

Because the amino acid sequence of the stem region of influenza virus is highly conserved, neutralizing antibodies elicited by proteins of the present invention may be broadly neutralizing. That is, neutralizing antibodies elicited by proteins of the present invention may neutralize influenza viruses of more than one type, subtype and/or strain. Thus, one embodiment of the present invention is a protein that elicits broadly neutralizing antibodies that bind the stem region of influenza HA protein. One embodiment is a protein that elicits antibodies that bind the stem region of an HA protein from more than one type of influenza virus selected from the group consisting of influenza type A viruses, influenza type B viruses and influenza type C viruses. One embodiment is a protein that elicits antibodies that bind the stem region of an HA protein from more than one sub-type of influenza virus selected from the group consisting of an H1 influenza virus, an H2 influenza virus, an influenza H3 virus, an influenza H4 virus, an influenza H5 virus, an influenza H6 virus, an H7 influenza virus, an H8 influenza virus, an H9 influenza virus, an H10 influenza virus, an H11 influenza virus, an H12 influenza virus, an H13 influenza virus, an H14 influenza virus, an H15 influenza virus and an H16 influenza virus. One embodiment is a protein that elicits antibodies that bind the stem region of an HA protein from more than one strain of influenza virus. One embodiment is a protein that elicits antibodies that bind the stem region of an HA protein from more than one strain of influenza virus selected from the viruses listed in Table 2. One embodiment of the present invention is a protein that elicits antibodies that bind more than one protein comprising an amino acid sequence at least 80% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. One embodiment of the present invention is a protein that elicits neutralizing antibodies that bind to more than one protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

Particularly useful proteins of the present invention are those comprising an immunogenic portion of an influenza HA protein. Thus, one embodiment of the present invention is a protein comprising at least one immunogenic portion from the stem region of influenza HA protein, wherein the protein elicits neutralizing antibodies against an influenza virus. Such a protein is referred to as a stem-region protein (or a stem-region immunogen). One embodiment of the present invention is a protein comprising at least one immunogenic portion from the stem region of an HA protein from a virus selected from the group consisting of influenza type A viruses, influenza type B viruses and influenza type C viruses, wherein the protein elicits neutralizing antibodies against an influenza virus. One embodiment of the present invention is a protein comprising at least one immunogenic portion from the stem region of an HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an influenza H3 virus HA protein, an influenza H4 virus HA protein, an influenza H5 virus HA protein, an influenza H6 virus HA protein, an H7 influenza virus HA protein, an H8 influenza virus HA protein, an H9 influenza virus HA protein, an H10 influenza

virus HA protein HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein and an H16 influenza virus HA protein. One embodiment of the present invention is a protein comprising at least one immunogenic portion from the stem region of an HA protein from the viruses listed in Table 2. One embodiment of the present invention is a protein comprising at least one immunogenic portion from a protein comprising an amino acid sequence at least 80% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. One embodiment of the present invention is a protein comprising at least one immunogenic portion from a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. In one embodiment, such proteins comprising immunogenic portions of the HA protein elicit the production of broadly neutralizing antibodies against influenza virus.

Immunogenic portions of proteins comprise epitopes, which are clusters of amino acid residues that are recognized by the immune system, thereby eliciting an immune response. Such epitopes may consist of contiguous amino acids residues (i.e., amino acid residues that are adjacent to one another in the protein), or they may consist of non-contiguous amino acid residues (i.e., amino acid residues that are not adjacent one another in the protein) but which are in close spatial proximity in the finally folded protein. It is well understood by those skilled in the art that epitopes require a minimum of six amino acid residues in order to be recognized by the immune system. Thus, in one embodiment the immunogenic portion from the influenza HA protein comprises at least one epitope. One embodiment of the present invention is a protein comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of influenza HA protein. One embodiment of the present invention is a protein comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of an HA protein from a virus selected from the group consisting of influenza type A viruses, influenza type B viruses and influenza type C viruses. One embodiment of the present invention is a protein comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of an HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an influenza H3 virus HA protein, an influenza H4 virus HA protein, an influenza H5 virus HA protein, an influenza H6 virus HA protein, an H7 influenza virus HA protein, an H8 influenza virus HA protein, an H9 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein and an H16 influenza virus HA protein. One embodiment of the present invention is a protein comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of an HA protein from a strain of virus selected

from the viruses listed in Table 2. In one embodiment, the amino acids are contiguous amino acids from the stem region of the HA protein. In one embodiment, such proteins comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of an HA protein elicit the production of broadly neutralizing antibodies against influenza virus. One embodiment of the present invention is a protein comprising at least 6 amino acids, at least 10 amino acids, at least 25 amino acids, at least 50 amino acids, at least 75 amino acids or at least 100 amino acids from the stem region of an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the amino acids are contiguous amino acids from the stem region of the HA protein. In one embodiment, the amino acids are non-contiguous, but are in close spatial proximity in the final protein.

While the present application discloses the use of stem regions from several exemplary HA proteins having specific sequences, the invention may also be practiced using stem regions from proteins comprising variations of the disclosed HA sequences. Thus, one embodiment of the present invention is a stem-region protein comprising an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical the stem region of an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. One embodiment of the present invention is a stem-region protein comprising an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. One embodiment of the present invention is a stem-region protein comprising the stem region of an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. One embodiment of the present invention is a stem-region protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

While the proteins disclosed thus far may elicit broadly neutralizing antibodies against an influenza virus, the inventors have discovered that such proteins are more stable and easier to purify when they exist in a trimeric form. Thus, one embodiment is a protein comprising the stem-region protein of the present invention joined to a trimerization domain. In one embodiment, the stem region is from an HA protein comprising an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID

NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region is from an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the stem region protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the trimerization domain is selected from the group consisting of the HIV-1 gp41 trimerization domain, the SIV gp41 trimerization domain, the Ebola virus gp-2 trimerization domain, the HTLV-1 gp-21 trimerization domain, the T4 fibritin trimerization domain (i.e., foldon), the yeast heat shock transcription factor trimerization domain, and the human collagen trimerization domain. In one embodiment, the trimerization domain is an HIV gp41 trimerization domain.

The inventors have also found that, in some instances, stem region proteins of the present invention may be more stable when joined to at least part of the head region of the HA protein. Thus, one embodiment of the present invention is a protein comprising a stem region protein joined to the head region of an HA protein and a trimerization domain. In one embodiment, the stem region protein is from an HA protein comprising an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region protein is from an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the stem region protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

In some embodiments of the present invention, the various protein domains (e.g., stem region protein, trimerization domain, head region, etc.) may be joined directly to one

another. In other embodiments, it may be necessary to employ linkers (also referred to as a spacer sequences) so that the various domains are in the proper special orientation. The linker sequence is designed to position the hemagglutinin protein in such a way to that it maintains the ability to elicit an immune response to the influenza virus. Linker sequences of the present invention comprise amino acids. Preferable amino acids to use are those having small side chains and/or those which are not charged. Such amino acids are less likely to interfere with proper folding and activity of the fusion protein. Accordingly, preferred amino acids to use in linker sequences, either alone or in combination are serine, glycine and alanine Examples of such linker sequences include, but are not limited to, SGG, GSG, GG and NGTGGSG. Amino acids can be added or subtracted as needed. Those skilled in the art are capable of determining appropriate linker sequences for proteins of the present invention.

One embodiment of the present invention is a fusion protein comprising a stem region protein joined to at least a portion of the head region of an HA protein and a trimerization domain, wherein the fusion protein comprises one or more linker sequences. In one embodiment, the stem region protein is from an HA protein comprising an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region protein is from an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the stem region protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the stem region protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the linker is selected from the group consisting of GG, GSG and NGTGGSG. In one embodiment, the protein elicits antibodies that neutralize at least one virus that is a different Type, sub-type or strain than the Type, sub-type or strain of the virus from which the HA protein was obtained.

Vaccines Comprising HA-Ferritin Fusion Proteins

The inventors have also discovered that fusion of influenza HA protein with ferritin protein (HA-ferritin fusion protein) results in a vaccine that elicits a robust immune response to influenza virus. Such HA-ferritin fusion proteins self-assemble into nanoparticles that display immunogenic portions of influenza hemagglutinin protein on their surface. These nanoparticles are useful for vaccinating individuals against a broad range of influenza viruses. Thus, one embodiment of the present invention is an HA-ferritin fusion protein comprising a monomeric ferritin subunit disclosed

31

herein joined to an influenza hemagglutinin protein disclosed herein, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles.

Ferritin is a globular protein found in all animals, bacteria, and plants, that acts primarily to control the rate and location of polynuclear $\text{Fe(III)}_2\text{O}_3$ formation through the transportation of hydrated iron ions and protons to and from a mineralized core. The globular form of ferritin is made up of monomeric subunit proteins (also referred to as monomeric ferritin subunits), which are polypeptides having a molecule weight of approximately 17-20 kDa. An example of the sequence of one such monomeric ferritin subunit is represented by SEQ ID NO:2. Each monomeric ferritin subunit has the topology of a helix bundle which includes a four antiparallel helix motif, with a fifth shorter helix (the C-terminal helix) lying roughly perpendicular to the long axis of the 4 helix bundle. According to convention, the helices are labeled 'A, B, C, and D & E' from the N-terminus respectively. The N-terminal sequence lies adjacent to the capsid three-fold axis and extends to the surface, while the E helices pack together at the four-fold axis with the C-terminus extending into the particle core. The consequence of this packing creates two pores on the capsid surface. It is expected that one or both of these pores represent the point by which the hydrated iron diffuses into and out of the capsid. Following production, these monomeric ferritin subunit proteins self-assemble into the globular ferritin protein. Thus, the globular form of ferritin comprises 24 monomeric, ferritin subunit proteins, and has a capsid-like structure having 432 symmetry.

According to the present invention, a monomeric ferritin subunit of the present invention is a full length, single polypeptide of a ferritin protein, or any portion thereof, which is capable of directing self-assembly of monomeric ferritin subunits into the globular form of the protein. Amino acid sequences from monomeric ferritin subunits of any known ferritin protein can be used to produce fusion proteins of the present invention, so long as the monomeric ferritin subunit is capable of self-assembling into a nanoparticle displaying hemagglutinin on its surface. In one embodiment, the monomeric subunit is from a ferritin protein selected from the group consisting of a bacterial ferritin protein, a plant ferritin protein, an algal ferritin protein, an insect ferritin protein, a fungal ferritin protein and a mammalian ferritin protein. In one embodiment, the ferritin protein is from *Helicobacter pylori*.

HA-ferritin fusion proteins of the present invention need not comprise the full-length sequence of a monomeric subunit polypeptide of a ferritin protein. Portions, or regions, of the monomeric ferritin subunit protein can be utilized so long as the portion comprises an amino acid sequence that directs self-assembly of monomeric ferritin subunits into the globular form of the protein. One example of such a region is located between amino acids 5 and 167 of the *Helicobacter pylori* ferritin protein. More specific regions are described in Zhang, Y. Self-Assembly in the Ferritin Nano-Cage Protein Super Family. 2011, Int. J. Mol. Sci., 12, 5406-5421, which is incorporated herein by reference in its entirety.

One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to at least 25 contiguous amino acids, at least 50 contiguous amino acids, at least 75 contiguous amino acids, at least 100 contiguous amino acids, or at least 150 contiguous amino acids from a monomeric ferritin subunit, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodi-

32

ment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to at least 25 contiguous amino acids, at least 50 contiguous amino acids, at least 75 contiguous amino acids, at least 100 contiguous amino acids, or at least 150 contiguous amino acids from the region of a ferritin protein corresponding to the amino acid sequences of the *Helicobacter pylori* ferritin monomeric subunit that direct self-assembly of the monomeric subunits into the globular form of the ferritin protein, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to at least 25 contiguous amino acids, at least 50 contiguous amino acids, at least 75 contiguous amino acids, at least 100 contiguous amino acids, or at least 150 contiguous amino acids from SEQ ID NO:2 that are capable of directing self-assembly of the monomeric subunits into the globular ferritin protein, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA-protein of the present invention joined to at least 25 contiguous amino acids, at least 50 contiguous amino acids, at least 75 contiguous amino acids, at least 100 contiguous amino acids, or at least 150 contiguous amino acids from amino acid residues 5-167 of SEQ ID NO:2, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to at least 25 contiguous amino acids, at least 50 contiguous amino acids, at least 75 contiguous amino acids, at least 100 contiguous amino acids, or at least 150 contiguous amino acids from SEQ ID NO:5, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to amino acid residues 5-167 from SEQ ID NO:2, or SEQ ID NO:5, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. As has been previously discussed, it is well-known in the art that some variations can be made in the amino acid sequence of a protein without affecting the activity of the protein. Such variations include insertion of amino acid residues, deletions of amino acid residues, and substitutions of amino acid residues. Thus, in one embodiment, the sequence of the monomeric ferritin subunit is divergent enough from the sequence of a ferritin subunit naturally found in a mammal, such that when the variant monomeric ferritin subunit is introduced into the mammal, it does not result in the production of antibodies that react with the mammal's natural ferritin protein. According to the present invention, such a monomeric subunit is referred to as immunogenically neutral. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, and at least 97% identical to the amino acid sequence of a monomeric ferritin subunit that is responsible for directing self-assembly of the monomeric ferritin subunits into the globular form of the protein, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. In one embodiment, the HA-ferritin fusion protein comprises a polypeptide sequence identical in sequence to a monomeric ferritin subunit. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to an amino acid sequence at

least 80%, at least 85%, at least 90%, at least 95%, and at least 97% identical to the amino acid sequence of a monomeric ferritin subunit from *Helicobacter pylori*, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, and at least 97% identical to a sequence selected from amino acid residues 5-167 from SEQ ID NO:2 and SEQ ID NO:5, wherein the HA-ferritin fusion protein is capable of self-assembling into nanoparticles. One embodiment of the present invention is an HA-ferritin fusion protein comprising an HA protein of the present invention joined to a sequence selected from amino acid residues 5-167 from SEQ ID NO:2 and SEQ ID NO:5.

In some embodiments, it may be useful to engineer mutations into the amino acid sequences of proteins of the present invention. For example, it may be useful to alter sites such as enzyme recognition sites or glycosylation sites in the monomeric ferritin subunit, the trimerization domain, or linker sequences, in order to give the fusion protein beneficial properties (e.g., solubility, half-life, mask portions of the protein from immune surveillance). In this regard, it is known that the monomeric subunit of ferritin is not glycosylated naturally. However, it can be glycosylated if it is expressed as a secreted protein in mammalian or yeast cells. Thus, in one embodiment, potential N-linked glycosylation sites in the amino acid sequences from the monomeric ferritin subunit are mutated so that the mutated ferritin subunit sequences are no longer glycosylated at the mutated site. One such sequence of a mutated monomeric ferritin subunit is represented by SEQ ID NO:5.

According to the present invention, the hemagglutinin protein portion of HA-ferritin fusion proteins of the present invention can be from any influenza virus, so long as the HA-ferritin fusion protein elicits an immune response against an influenza virus. Thus, one embodiment of the preset invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence from an HA protein from an influenza A virus, an influenza B virus or an influenza C virus. One embodiment of the preset invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence from an influenza A Group 1 virus HA protein. One embodiment of the preset invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence from an influenza A Group 2 virus HA protein. One embodiment of the preset invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence from an HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein. One embodiment of the preset invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence from an HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein, and an H15 influenza virus HA protein. One embodiment of the preset invention is an

HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence of an HA protein from a virus listed in Table 2.

Preferred hemagglutinin proteins to use in constructing HA-ferritin fusion proteins of the present invention are those that elicit an immune response against an influenza virus. Even more preferred hemagglutinin proteins are those that are capable of eliciting antibodies to an influenza virus. One embodiment of the present invention is an HA-ferritin fusion protein that elicits antibodies to a virus selected from the group consisting of influenza A viruses, influenza B viruses and influenza C viruses. One embodiment of the present invention is a HA-ferritin fusion protein that elicits antibodies to a virus selected from the group consisting of an H1 influenza virus, an H2 influenza virus, an influenza H3 virus, an influenza H4 virus, an influenza H5 virus, an influenza H6 virus, an H7 influenza virus, an H8 influenza virus, an H9 influenza virus, an H10 influenza virus, an H11 influenza virus, an H12 influenza virus, an H13 influenza virus, an H14 influenza virus, an H15 influenza virus and an H16 influenza virus. One embodiment of the present invention is an HA-ferritin fusion protein that elicits antibodies to a virus listed in Table 2. Preferred antibodies elicited by HA-ferritin fusion proteins of the present invention are those that neutralize an influenza virus. Thus, one embodiment of the present invention is an HA-ferritin fusion protein that elicits neutralizing antibodies to a virus selected from the group consisting of influenza A viruses, influenza B viruses and influenza C viruses. One embodiment of the present invention is an HA-ferritin fusion protein that elicits neutralizing antibodies to a virus having a subtype selected from the group consisting of an H1 influenza virus, an H2 influenza virus, an influenza H3 virus, an influenza H4 virus, an influenza H5 virus, an influenza H6 virus, an H7 influenza virus, an H8 influenza virus, an H9 influenza virus, an H10 influenza virus, an H11 influenza virus, an H12 influenza virus, an H13 influenza virus, an H14 influenza virus, an H15 influenza virus and an H16 influenza virus. One embodiment of the present invention is an HA-ferritin fusion protein that elicits neutralizing antibodies to a virus listed in Table 2.

As has been discussed, neutralizing antibodies elicited by a HA-ferritin fusion protein of the present invention can neutralize viral infections by affecting any step in the life cycle of the virus. Thus, in one embodiment of the present invention, an HA-ferritin fusion protein elicits neutralizing antibodies that prevent influenza virus from attaching to the host cell. In one embodiment of the present invention, an HA-ferritin fusion protein may elicit neutralizing antibodies that prevent influenza virus from entering the host cell. In one embodiment of the present invention, an HA-ferritin fusion protein may elicit neutralizing antibodies that prevent fusion of viral membranes with endosomal membranes. In one embodiment of the present invention, an HA-ferritin fusion protein may elicit neutralizing antibodies that prevent influenza virus from releasing ribonucleoproteins into the cytoplasm of the host cell. In one embodiment of the present invention, an HA-ferritin fusion protein may elicit neutralizing antibodies that prevent assembly of new virus in the infected host cell. In one embodiment of the present invention, an HA-ferritin fusion protein may elicit neutralizing antibodies that prevent release of newly formed virus from the infected host cell.

Preferred HA-ferritin fusion proteins of the present invention are those that elicit broadly neutralizing antibodies. Thus, one embodiment is an HA-ferritin fusion protein that elicits antibodies that neutralizes more than one type of

influenza virus selected from the group consisting of influenza type A viruses, influenza type B viruses and influenza type C viruses. One embodiment is an HA-ferritin fusion protein that elicits antibodies that neutralize more than one sub-type of influenza virus selected from the group consisting of an H1 influenza virus, an H2 influenza virus, an influenza H3 virus, an influenza H4 virus, an influenza H5 virus, an influenza H6 virus, an H7 influenza virus, an H8 influenza virus, an H9 influenza virus, an H10 influenza virus, an H11 influenza virus, an H12 influenza virus, an H13 influenza virus, an H14 influenza virus, an H15 influenza virus and an H16 influenza virus. One embodiment is an HA-ferritin protein that elicits antibodies that neutralize from more than one strain of influenza virus selected from the viruses listed in Table 2.

It will be understood by those skilled in the art that particularly useful HA-ferritin useful proteins of the present invention are those comprising an immunogenic portion of influenza HA protein. Thus, one embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least one immunogenic portion of an influenza HA protein. One embodiment of the present invention is an HA-ferritin protein comprising a ferritin protein of the present invention joined to at least one immunogenic portion of an HA protein from a virus selected from the group consisting of influenza type A viruses, influenza type B viruses and influenza type C viruses. One embodiment of the present invention is an HA-ferritin protein comprising a ferritin protein of the present invention joined to at least one immunogenic portion of an HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least one immunogenic portion of an HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein, and an H16 influenza virus HA protein, joined to a ferritin protein of the present invention. One embodiment of the present invention is an HA-ferritin protein comprising a ferritin protein of the present invention joined to at least one immunogenic portion of an HA protein from virus listed in Table 2. In one embodiment, an HA-ferritin fusion protein comprising an immunogenic portion of an HA protein elicits the production of broadly neutralizing antibodies against influenza virus.

Immunogenic portions of proteins comprise epitopes, which are clusters of amino acid residues that are recognized by the immune system, thus eliciting an immune response. Such epitopes may consist of contiguous amino acids residues (i.e., amino acid residues that are adjacent to one another in the protein), or they may consist of non-contiguous amino acid residues (i.e., amino acid residues that are not adjacent one another in the protein) but which are in close spatial proximity in the finally folded protein. It is well understood by those skilled in the art that such epitopes require a minimum of six amino acid residues in order to be recognized by the immune system. Thus, one embodiment of the present invention is an HA-ferritin fusion comprising

an immunogenic portion from the influenza HA protein, wherein the immunogenic portion comprises at least one epitope.

It is known in the art that some variation in a protein sequence can be tolerated without significantly affecting the activity of the protein. Thus, one embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence that is a variant of an HA protein from a virus selected from the group consisting of influenza Type A viruses influenza Type B viruses and influenza type C viruses. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of a HA protein from a virus selected from the group consisting of influenza Type A viruses influenza Type B viruses and influenza type C viruses, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of a HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of a HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein, and an H16 influenza virus HA protein, joined to a ferritin protein of the present invention, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of a HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein, and an H16 influenza virus HA protein, joined to a ferritin protein of the present invention, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of a HA protein from a virus listed in Table 2, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza

protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38.

One embodiment of the present invention is an HA-ferritin fusion protein comprising an amino acid sequence at least 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to a sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. One embodiment of the present invention is an HA-ferritin fusion protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68.

It is known in the art that influenza hemagglutinin proteins have various regions, or domains, each possessing specific activities. For example, the globular head extends out from the lipid membrane and comprises two domains: the receptor binding domain (RBD) and the vestigial esterase domain. The RB domain is involved in binding of the HA protein to receptors. The globular head also includes several antigenic sites that include immunodominant epitopes. The stem region is responsible for anchoring the HA protein into the viral lipid envelope. Thus, it will be understood by those skilled in the art that HA-ferritin fusion proteins of the present invention need not comprise the entire sequence of the HA protein. Instead, an HA-ferritin fusion protein can comprise only those portions, regions, domains, and the like, that contain the necessary activities for practicing the present invention. For example, an HA-ferritin fusion protein may contain only those amino acid sequences from the HA protein that contain antigenic sites, epitopes, immunodominant epitopes, and the like.

One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from an HA protein from a virus selected from the group consisting of influenza Type A viruses, influenza Type B viruses and influenza type C viruses, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from an HA protein selected from the group consisting of an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150

amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from an HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, an H15 influenza virus HA protein, and an H16 influenza virus HA protein, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from and HA protein from a virus listed in Table 2, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against in influenza virus. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a protein consisting of a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38.

One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least one domain from a HA protein from a virus listed in Table 2, wherein the domain is selected from the group consisting of an ectodomain, an RDB domain, a stem domain, and a domain comprising the region stretching from the amino acid residue immediately distal to the last amino acid of second helical coil to the amino acid residue proximal to the first amino acid of the transmembrane domain. According to the present invention, an ectodomain of an influenza hemagglutinin protein refers to the portion of the hemagglutinin protein that lies outside its transmembrane domain. In one embodiment, the HA-ferritin fusion protein comprises a ferritin protein of the present invention joined to a region of a HA protein from a virus listed in Table 2, wherein the region consists of the amino acid immediately distal to the last amino acid of the second helical coiled coil and proximal to the first amino acid of the transmembrane domain. In one embodiment, the HA-ferritin fusion protein comprises a ferritin protein of the present invention joined to a region of a HA protein from a virus listed in Table 2, wherein the region comprises an amino acid sequence distal to the second helical coiled coil and proximal to the transmembrane domain. In one embodiment, the HA-ferritin fusion protein comprises a ferritin protein of the present invention joined to the ectodomain of a HA protein from a virus listed in Table 2. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38.

The stem region of an influenza HA protein is a particularly useful domain for constructing fusion proteins of the present invention. Thus, one embodiment of the present invention is a ferritin protein of the present invention joined to at least one immunogenic portion from the stem region of influenza HA protein. According to the preset invention, such a protein is referred to an HA SS-ferritin fusion protein. As used herein, the HA stem region of the hemagglutinin protein consists of the amino acids from the membrane up to the head region of the protein. More specifically, the stem region consists of the amino terminal amino acid up to the cysteine at position 52, and all residues after the cysteine residue at position 277 (using standard H3 numbering). Sequences of exemplary stem regions are represented by SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region of an HA protein from a virus selected from the group consisting of influenza Type A viruses influenza Type B viruses and influenza type C viruses, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region of an HA protein selected from the group consisting an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region of an HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, and an H16 influenza virus HA protein, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region of an HA protein from a virus listed in Table 2, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza virus. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region of an HA protein comprising a sequence selected

from the group consisting of SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from the stem region comprising a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of the stem region of an HA protein from a virus selected from the group consisting of influenza Type A viruses influenza Type B viruses and influenza type C viruses, wherein the Ha-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of the stem region of an HA protein selected from the group consisting an H1 influenza virus HA protein, an H2 influenza virus HA protein, an H5 influenza virus HA protein, an H7 virus influenza HA protein and an H9 influenza virus HA protein, wherein the Ha-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of the stem region of an HA protein selected from the group consisting of an H3 influenza virus HA protein, an H4 influenza virus HA protein, an H6 influenza virus HA protein, an H8 influenza virus HA protein, an H10 influenza virus HA protein, an H11 influenza virus HA protein, an H12 influenza virus HA protein, an H13 influenza virus HA protein, an H14 influenza virus HA protein, and an H16 influenza virus HA protein, joined to a ferritin protein of the present invention, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the sequence of the stem region of an HA protein from a virus listed in Table 2, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to the stem region of an HA protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23,

41

SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising a ferritin protein of the present invention joined to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

As has been described for stem region proteins of the present invention, the inventors have found that HA-ferritin fusion proteins are more stable and easier to purify when they exist in a trimeric form. Thus, in one embodiment of the present invention the HA portion of the HA-ferritin fusion protein is joined to one or more trimerization domains. In one embodiment, the HA protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38, joined to one or more trimerization domains. In one embodiment, the HA protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38 joined to one or more trimerization domains. In one embodiment, the HA protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 92% identical, at least 94% identical, at least 96% identical, at least 98% identical or at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98 joined to one or more trimerization domains. In one embodiment, the HA protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98 joined to one or more trimerization domains. In one embodiment, the trimerization domain is selected from the group consisting of the HIV-1 gp41 trimerization domain, the SIV gp41 trimerization domain, the Ebola virus gp-2 trimerization domain, the HTLV-1 gp-21 trimerization domain, the T4 fibritin trimerization domain (i.e., foldon), the yeast heat shock transcription factor trimerization domain, and the human collagen trimerization domain. In one embodiment, the trimerization domain is an HIV gp41 trimerization domain.

42

Additionally, the inventors have found that, in some instances, HA-ferritin fusion proteins in which the HA portion is limited to HA stem region sequences may be more stable when joined to at least part of the head region of the HA protein. Thus, one embodiment of the present invention is an HA SS-ferritin fusion protein, wherein, the HA portion of the fusion protein is joined to an amino acid sequence from at least a portion of an HA protein head region.

HA-ferritin proteins of the present invention are constructed by joining ferritin proteins of the present invention with HA proteins of the present invention. In addition, HA-ferritin fusion proteins may contain other domains (e.g., stem region protein, trimerization domain, head region, etc.) that improve the functionality of the final HA-ferritin fusion protein. In some embodiments, joining of the various proteins and/or domains can be done such that the sequences are directly linked. In other embodiments, it may be necessary to employ linkers (also referred to as a spacer sequences) between the various proteins and/or domains so that the so that they are in the proper special orientation. More specifically, linker sequence can be inserted so that the hemagglutinin protein is positioned in such a way to maintain the ability to elicit an immune response to the influenza virus. Linker sequences of the present invention comprise amino acids. Preferable amino acids to use are those having small side chains and/or those which are not charged. Such amino acids are less likely to interfere with proper folding and activity of the fusion protein. Accordingly, preferred amino acids to use in linker sequences, either alone or in combination are serine, glycine and alanine Examples of such linker sequences include, but are not limited to, SGG, GSG, GG and NGTGGSG. Amino acids can be added or subtracted as needed. Those skilled in the art are capable of determining appropriate linker sequences for proteins of the present invention.

In accordance with the invention, suitable portions of the hemagglutinin protein can be joined to the ferritin protein either as an exocapsid product by fusion with the N-terminal sequence lying adjacent to the capsid three-fold axis, as an endocapsid product by fusion with the C-terminus extending inside the capsid core, or a combination thereof. In one embodiment, the hemagglutinin portion of the fusion protein is joined to the N-terminal sequence of the ferritin portion of the fusion protein.

One embodiment of the present invention is an HA-ferritin fusion protein comprising an amino acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128, wherein the HA-ferritin fusion protein elicits the production of neutralizing antibodies against an influenza protein. One embodiment of the present invention is an HA-ferritin fusion protein comprising SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

Proteins of the present invention are encoded by nucleic acid molecules of the present invention. In addition, they are expressed by nucleic acid constructs of the present invention. As used herein a nucleic acid construct is a recombinant expression vector, i.e., a vector linked to a nucleic acid molecule encoding a protein such that the nucleic acid molecule can effect expression of the protein when the nucleic acid construct is administered to, for example, a subject or an organ, tissue or cell. The vector also enables

transport of the nucleic acid molecule to a cell within an environment, such as, but not limited to, an organism, tissue, or cell culture. A nucleic acid construct of the present disclosure is produced by human intervention. The nucleic acid construct can be DNA, RNA or variants thereof. The vector can be a DNA plasmid, a viral vector, or other vector. In one embodiment, a vector can be a cytomegalovirus (CMV), retrovirus, adenovirus, adeno-associated virus, herpes virus, vaccinia virus, poliovirus, sindbis virus, or any other DNA or RNA virus vector. In one embodiment, a vector can be a pseudotyped lentiviral or retroviral vector. In one embodiment, a vector can be a DNA plasmid. In one embodiment, a vector can be a DNA plasmid comprising viral components and plasmid components to enable nucleic acid molecule delivery and expression. Methods for the construction of nucleic acid constructs of the present disclosure are well known. See, for example, *Molecular Cloning: a Laboratory Manual*, 3rd edition, Sambrook et al. 2001 Cold Spring Harbor Laboratory Press, and *Current Protocols in Molecular Biology*, Ausubel et al. eds., John Wiley & Sons, 1994. In one embodiment, the vector is a DNA plasmid, such as a CMV/R plasmid such as CMV/R or CMV/R 8 KB (also referred to herein as CMV/R 8 kb). Examples of CMV/R and CMV/R 8 kb are provided herein. CMV/R is also described in U.S. Pat. No. 7,094,598 B2, issued Aug. 22, 2006.

As used herein, a nucleic acid molecule comprises a nucleic acid sequence that encodes a stem region immunogen, a ferritin monomeric subunit, a hemagglutinin protein, and/or an HA-ferritin fusion protein of the present invention. A nucleic acid molecule can be produced recombinantly, synthetically, or by a combination of recombinant and synthetic procedures. A nucleic acid molecule of the disclosure can have a wild-type nucleic acid sequence or a codon-modified nucleic acid sequence to, for example, incorporate codons better recognized by the human translation system. In one embodiment, a nucleic acid molecule can be genetically-engineered to introduce, or eliminate, codons encoding different amino acids, such as to introduce codons that encode an N-linked glycosylation site. Methods to produce nucleic acid molecules of the disclosure are known in the art, particularly once the nucleic acid sequence is known. It is to be appreciated that a nucleic acid construct can comprise one nucleic acid molecule or more than one nucleic acid molecule. It is also to be appreciated that a nucleic acid molecule can encode one protein or more than one protein.

Preferred nucleic acid molecules are those that encode a stem-region protein, a ferritin monomeric subunit, a hemagglutinin protein, and/or an HA-ferritin fusion protein comprising a monomeric subunit of a ferritin protein joined to an influenza hemagglutinin protein. Thus, one embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence encoding a protein that comprises a monomeric subunit of a ferritin protein joined to an influenza hemagglutinin protein. In one embodiment, the monomeric subunit of ferritin is from the ferritin protein of *Helicobacter pylori*. In one embodiment, the monomeric subunit comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5. In one embodiment, the monomeric subunit comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5. In one embodiment the influenza hemagglutinin protein comprises an amino acid sequence at least 80%, at least 85%, at least

90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment the influenza hemagglutinin protein comprises a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment the influenza hemagglutinin protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment the influenza hemagglutinin protein comprises a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment the influenza hemagglutinin protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment the influenza hemagglutinin protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, or at least 200 amino acids from a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic sequence encoding a protein comprising an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic sequence encoding a protein comprising an amino acid sequence at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 96%, at least 98% or at least 99% identical to SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID

45

NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

Also embodied in the present invention are nucleic acid sequences that are variants of nucleic acid sequence encoding protein of the present invention. Such variants include nucleotide insertions, deletions, and substitutions, so long as they do not affect the ability of fusion proteins of the present invention to self-assemble into nanoparticles, or significantly affect the ability of the hemagglutinin portion of fusion proteins to elicit an immune response to an influenza virus. Thus, one embodiment of the present invention is a nucleic acid molecule encoding a fusion protein of the present invention, wherein the monomeric subunit is encoded by a nucleotide sequence at least 85%, at least 90%, at least 95%, or at least 97% identical to a nucleotide sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:4. One embodiment of the present invention is a nucleic acid molecule encoding an HA-ferritin fusion protein of the present invention, wherein the HA protein is encoded by a nucleotide sequence at least 85%, at least 90%, at least 95%, at least 97% identical or at least 99% identical to a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. One embodiment of the present invention is a nucleic acid molecule encoding an HA-ferritin fusion protein of the present invention, wherein the HA protein is encoded by a nucleotide sequence at least 85%, at least 90%, at least 95%, at least 97% identical or at least 99% identical to a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:7, SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:19, SEQ ID NO:22, SEQ ID NO:25, SEQ ID NO:28, SEQ ID NO:31, SEQ ID NO:34, and SEQ ID NO:37. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:7, SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:19, SEQ ID NO:22, SEQ ID NO:25, SEQ ID NO:28, SEQ ID NO:31, SEQ ID NO:34, and SEQ ID NO:37.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:40, SEQ ID NO:43, SEQ ID NO:46, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:58, SEQ ID NO:61, SEQ ID NO:64, and SEQ ID NO:67. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:40, SEQ ID NO:43, SEQ ID NO:46, SEQ ID NO:49, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:58, SEQ ID NO:61, SEQ ID NO:64, and SEQ ID NO:67.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence at least about 80%, at least about 85%, at least about 90%, at least about

46

95%, at least about 97% or at least about 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:70, SEQ ID NO:73, SEQ ID NO:76, SEQ ID NO:79, SEQ ID NO:82, SEQ ID NO:85, SEQ ID NO:88, SEQ ID NO:91, SEQ ID NO:94, and SEQ ID NO:97. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:70, SEQ ID NO:73, SEQ ID NO:76, SEQ ID NO:79, SEQ ID NO:82, SEQ ID NO:85, SEQ ID NO:88, SEQ ID NO:91, SEQ ID NO:94, and SEQ ID NO:97.

One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:100, SEQ ID NO:103, SEQ ID NO:106, SEQ ID NO:109, SEQ ID NO:112, SEQ ID NO:115, SEQ ID NO:121, SEQ ID NO:124, and SEQ ID NO:127. One embodiment of the present invention is a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:100, SEQ ID NO:103, SEQ ID NO:106, SEQ ID NO:109, SEQ ID NO:112, SEQ ID NO:115, SEQ ID NO:121, SEQ ID NO:124, and SEQ ID NO:127.

Also encompassed by the present invention are expression systems for producing fusion proteins of the present invention. In one embodiment, nucleic acid molecules of the present invention are operationally linked to a promoter. As used herein, operationally linked means that proteins encoded by the linked nucleic acid molecules can be expressed when the linked promoter is activated. Promoters useful for practicing the present invention are known to those skilled in the art. One embodiment of the present invention is a recombinant cell comprising a nucleic acid molecule of the present invention. One embodiment of the present invention is a recombinant virus comprising a nucleic acid molecule of the present invention.

As indicated above, the recombinant production of the ferritin fusion proteins of the present invention can take place using any suitable conventional recombinant technology currently known in the field. For example, molecular cloning a fusion protein, such as ferritin with a suitable protein such as the recombinant influenza hemagglutinin protein, can be carried out via expression in *E. coli* with the suitable monomeric subunit protein, such as the *helicobacter pylori* ferritin monomeric subunit. The construct may then be transformed into protein expression cells, grown to suitable size, and induced to produce the fusion protein.

As has been described, because HA-ferritin fusion proteins of the present invention comprise a monomeric subunit of ferritin, they can self-assemble. According to the present invention, the supramolecule resulting from such self-assembly is referred to as a hemagglutinin expressing ferritin based nanoparticle. For ease of discussion, the hemagglutinin expressing ferritin based nanoparticle will simply be referred to as a, or the, nanoparticle (np). Nanoparticles of the present invention have the same structural characteristics as the ferritin proteins described earlier. That is, they contain 24 subunits and have 432 symmetry. In the case of nanoparticles of the present invention, the subunits are the fusion proteins comprising a ferritin monomeric subunit joined to an influenza hemagglutinin protein. Such nanoparticles display at least a portion of the hemagglutinin protein on their surface as hemagglutinin trimers. In such a construction, the hemagglutinin trimer is accessible to the immune system and thus can elicit an immune response. Thus, one embodiment

ment of the present invention is a nanoparticle comprising an HA-ferritin fusion protein, wherein the fusion protein comprises a monomeric ferritin subunit joined to an influenza hemagglutinin protein. In one embodiment, the nanoparticle is an octahedron. In one embodiment, the nanoparticle displays the hemagglutinin protein on its surface as a hemagglutinin trimer. In one embodiment, the influenza hemagglutinin protein is capable of eliciting neutralizing antibodies to an influenza virus. In one embodiment, the monomeric ferritin subunit comprises at least 50 amino acids, at least 100 amino acids, or at least 150 amino acids from an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5, and/or comprises an amino acid sequence at least 85%, at least 90%, at least 95%, at least 97% at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5. In one embodiment, the monomeric ferritin subunit comprises SEQ ID NO:2 or SEQ ID NO:5.

In one embodiment, the influenza hemagglutinin protein comprises at least one epitope from an influenza hemagglutinin protein listed in Table 2. In one embodiment, the influenza hemagglutinin protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a hemagglutinin protein of a virus listed in Table 2. In one embodiment, the hemagglutinin protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a protein consisting of a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the hemagglutinin protein comprises a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38.

In one embodiment, the influenza hemagglutinin protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to the sequence of an hemagglutinin protein from a virus listed in Table 2. In one embodiment, the influenza hemagglutinin protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a protein sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38.

In one embodiment, the hemagglutinin protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a protein consisting of a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the influenza hemagglutinin protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a protein sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the hemagglutinin protein comprises a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98.

In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a protein sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128. In one embodiment, the HA-ferritin fusion protein comprises SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

Because stem region immunogens, HA-ferritin fusion proteins and nanoparticles of the present invention can elicit an immune response to an influenza virus, they can be used as vaccines to protect individuals against infection by influenza virus. According to the present invention a vaccine can be a stem region immunogen, an HA-ferritin fusion protein, or a nanoparticle of the present invention. Thus, one embodiment of the present invention is a vaccine comprising a stem region immunogen, an HA-ferritin fusion protein, or a nanoparticle of the present invention. Vaccines of the present invention can also contain other components such as adjuvants, buffers and the like. Although any adjuvant can be used, preferred embodiments can contain: chemical adjuvants such as aluminum phosphate, benzylalkonium chloride, ubenimex, and QS21; genetic adjuvants such as the IL-2 gene or fragments thereof, the granulocyte macrophage colony-stimulating factor (GM-CSF) gene or fragments thereof, the IL-18 gene or fragments thereof, the chemokine (C—C motif) ligand 21 (CCL21) gene or fragments thereof, the IL-6 gene or fragments thereof, CpG, LPS, TLR agonists, and other immune stimulatory genes; protein adjuvants such IL-2 or fragments thereof, the granulocyte macrophage colony-stimulating factor (GM-CSF) or fragments thereof, IL-18 or fragments thereof, the chemokine (C—C motif) ligand 21 (CCL21) or fragments thereof, IL-6 or fragments thereof, CpG, LPS, TLR agonists and other immune stimulatory cytokines or fragments thereof; lipid adjuvants such as cationic liposomes, N3 (cationic lipid), monophosphoryl lipid A (MPL1); other adjuvants including cholera toxin, enterotoxin, Fms-like tyrosine kinase-3 ligand (Flt-3L), bupivacaine, marcaine, and levamisole.

One embodiment of the disclosure is a ferritin-based nanoparticle vaccine that includes more than one influenza hemagglutinin protein. Such a vaccine can include a combination of different influenza hemagglutinin proteins, either on a single nanoparticle or as a mixture of nanoparticles, at least two of which have a unique influenza hemagglutinin

proteins. A multivalent vaccine can comprise as many influenza hemagglutinin proteins as necessary in order to result in production of the immune response necessary to protect against a desired breadth of virus strains. In one embodiment, the vaccine comprises a hemagglutinin protein from at least two different influenza strains (bi-valent). In one embodiment, the vaccine comprises a hemagglutinin protein from at least three different influenza strains (tri-valent). In one embodiment, the vaccine comprises a hemagglutinin protein from at least four different influenza strains (tetra-valent). In one embodiment, the vaccine comprises a hemagglutinin protein from at least five different influenza strains (penta-valent). In one embodiment, the vaccine comprises a hemagglutinin protein from at least six different influenza strains (hexa-valent). In various embodiments, a vaccine comprises a hemagglutinin protein from each of 7, 8, 9, or 10 different strains of influenza virus. An example of such a combination is a ferritin-based nanoparticle vaccine that comprises influenza A group 1 hemagglutinin protein, an influenza A group 2 hemagglutinin protein, and an influenza B hemagglutinin protein. In one embodiment, the influenza hemagglutinin proteins are H1 HA, H3 HA, and B HA. In one embodiment, the influenza hemagglutinin proteins are those included in the 2011-2012 influenza vaccine. Another example of a multivalent vaccine is a ferritin based nanoparticle vaccine that comprises hemagglutinin proteins from four different influenza viruses. In one embodiment, the multivalent vaccine comprises hemagglutinin proteins from H1 A/NC/20/1999, H1 A/CA/04/2009, H2 A/Singapore/1/1957 and H5 A/Indonesia/05/2005. Such a vaccine is described in Example 2.

One embodiment of the present invention is a method to vaccinate an individual against influenza virus, the method comprising administering a nanoparticle to an individual such that an immune response against influenza virus is produced in the individual, wherein the nanoparticle comprises a monomeric subunit from ferritin joined to an influenza hemagglutinin protein, and wherein the nanoparticle displays the influenza hemagglutinin on its surface. In one embodiment, the nanoparticle is a monovalent nanoparticle. In one embodiment, the nanoparticle is multivalent nanoparticle. Another embodiment of the present invention is a method to vaccinate an individual against infection with influenza virus, the method comprising:

a) obtaining a nanoparticle comprising monomeric subunits, wherein the monomeric subunits comprise a ferritin protein joined to an influenza hemagglutinin protein, and wherein the nanoparticle displays the influenza hemagglutinin on its surface; and,

b) administering the nanoparticle to an individual such that an immune response against an influenza virus is produced.

One embodiment of the present invention is a method to vaccinate an individual against influenza virus, the method comprising administering a vaccine of the embodiments to an individual such that an immune response against influenza virus is produced in the individual, wherein the vaccine comprises at least one nanoparticle comprising a monomeric subunit from ferritin joined to an influenza hemagglutinin protein, and wherein the nanoparticle displays the influenza hemagglutinin on its surface. In one embodiment, the vaccine is a stem region immunogen. In one embodiment, the vaccine is a nanoparticle. In one embodiment, the vaccine is a monovalent vaccine. In one embodiment, the vaccine is multivalent vaccine. Another embodiment of the present

invention is a method to vaccinate an individual against infection with influenza virus, the method comprising:

a) obtaining a vaccine comprising at least one nanoparticle comprising an HA-ferritin fusion protein, wherein the fusion protein comprises a ferritin protein joined to an influenza HA protein, and wherein the nanoparticle displays the influenza HA on its surface; and,

b) administering the vaccine to an individual such that an immune response against an influenza virus is produced.

In one embodiment, the nanoparticle is a monovalent nanoparticle. In one embodiment, the nanoparticle is multivalent nanoparticle.

In one embodiment, the nanoparticle is an octahedron. In one embodiment, the influenza hemagglutinin protein is capable of eliciting neutralizing antibodies to an influenza virus. In one embodiment, the influenza HA protein is capable of eliciting broadly neutralizing antibodies to an influenza virus. In one embodiment, the ferritin portion of the fusion protein comprise at least 50 amino acids, at least 100 amino acids, or at least 150 amino acids from an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5, and/or comprises an amino acid sequence at least 85%, at least 90%, at least 95%, at least 97% at least 99% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:5. In one embodiment, the HA portion of the fusion protein comprises at least one epitope from an influenza hemagglutinin protein listed in Table 2. In one embodiment, the HA portion of the fusion protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a hemagglutinin protein of a virus listed in Table 2. In one embodiment, the HA portion of the fusion protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a protein consisting of a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the HA portion of the fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to the sequence of an HA protein from a virus listed in Table 2. In one embodiment, the HA portion of the fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID NO:38. In one embodiment, the HA portion of the fusion protein comprises at least 25 amino acids, at least 50 amino acids, at least 75 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 300 amino acids, at least 400 amino acids, or at least 500 amino acids from a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the HA portion of the fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a

sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95, and SEQ ID NO:98. In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to a protein sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. In one embodiment, the HA-ferritin fusion protein comprises an amino acid sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% identical to SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128. In one embodiment, the HA-ferritin fusion protein comprises SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

Vaccines of the present invention can be used to vaccinate individuals using a prime/boost protocol. Such a protocol is described in U.S. Patent Publication No. 20110177122, which is incorporated herein by reference in its entirety. In such a protocol, a first vaccine composition may be administered to the individual (prime) and then after a period of time, a second vaccine composition may be administered to the individual (boost). Administration of the boosting composition is generally weeks or months after administration of the priming composition, preferably about 2-3 weeks or 4 weeks, or 8 weeks, or 16 weeks, or 20 weeks, or 24 weeks, or 28 weeks, or 32 weeks. In one embodiment, the boosting composition is formulated for administration about 1 week, or 2 weeks, or 3 weeks, or 4 weeks, or 5 weeks, or 6 weeks, or 7 weeks, or 8 weeks, or 9 weeks, or 16 weeks, or 20 weeks, or 24 weeks, or 28 weeks, or 32 weeks after administration of the priming composition.

The first and second vaccine compositions can be, but need not be, the same composition. Thus, in one embodiment of the present invention, the step of administering the vaccine comprises administering a first vaccine composition, and then at a later time, administering a second vaccine composition. In one embodiment, the first vaccine composition comprises a nanoparticle comprising an HA-ferritin fusion protein of the present invention. In one embodiment, the first vaccine composition comprises a nanoparticle comprising an ectodomain from the hemagglutinin protein of an influenza virus selected from the group consisting of A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), B/Brisbane/60/2008 (2008 Bris, B). In one embodiment, the hemagglutinin of the first vaccine composition comprises an amino acid sequence at least about 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, and SEQ ID

NO:38. In one embodiment, the first vaccine composition comprises an HA-ferritin fusion protein comprising an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 97% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68, wherein the nanoparticle elicits an immune response against an influenza virus. In one embodiment, the first vaccine composition comprises an HA-ferritin fusion protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, and SEQ ID NO:68. In one embodiment, second vaccine composition comprises a nanoparticle comprising an HA SS-ferritin fusion protein of the present invention. In one embodiment, the HA SS-ferritin fusion protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 97% identical or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. In one embodiment, the HA SS-ferritin fusion protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98. In one embodiment, the HA SS-ferritin fusion protein comprises an amino acid sequence at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 97% identical or at least 99% identical to SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128, wherein the HA SS-ferritin fusion protein elicits an immune response to an influenza virus. In one embodiment, the HA SS-ferritin fusion protein comprises SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128. In one embodiment, the individual is at risk for infection with influenza virus. In one embodiment, the individual has been exposed to influenza virus. As used herein, the terms exposed, exposure, and the like, indicate the subject has come in contact with a person or animal that is known to be infected with an influenza virus. Vaccines of the present invention may be administered using techniques well known to those in the art. Techniques for formulation and administration may be found, for example, in "Remington's Pharmaceutical Sciences", 18th ed., 1990, Mack Publishing Co., Easton, Pa. Vaccines may be administered by means including, but not limited to, traditional syringes, needleless injection devices, or microprojectile bombardment gene guns. Suitable routes of administration include, but are not limited to, parenteral delivery, such as intramuscular, intradermal, subcutaneous, intramedullary injections, as well as, intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections, just to name a few. For injection, the compounds of one embodiment of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer.

In one embodiment, vaccines, or nanoparticles, of the present invention can be used to protect an individual against infection by heterologous influenza virus. That is, a vaccine made using hemagglutinin protein from one strain of influenza virus is capable of protecting an individual against infection by different strains of influenza. For example, a vaccine made using hemagglutinin protein from influenza A/New Caledonia/20/1999 (1999 NC, H1), can be used to protect an individual against infection by an influenza virus including, but not limited to A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 indo, H5), A/Perth/16/2009 (2009 Per, H3), and/or A/Brisbane/59/2007 (2007 Bris, H1).

In one embodiment, vaccines, or nanoparticles, of the present invention can be used to protect an individual against infection by an antigenically divergent influenza virus. Antigenically divergent refers to the tendency of a strain of influenza virus to mutate over time, thereby changing the amino acids that are displayed to the immune system. Such mutation over time is also referred to as antigenic drift. Thus, for example, a vaccine made using hemagglutinin protein from a A/New Caledonia/20/1999 (1999 NC, H1) strain of influenza virus is capable of protecting an individual against infection by earlier, antigenically divergent New Caledonia strains of influenza, and by evolving (or diverging) influenza strains of the future.

EXAMPLES

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the embodiments, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, and temperature is in degrees Celsius. Standard abbreviations are used.

Example 1

Design and Production of Ferritin-Based Nanoparticles Expressing Influenza Virus HA

This Example demonstrates the ability of HA-ferritin fusion proteins to form nanoparticles. Analysis of ferritin structure suggested that it was possible to insert a heterologous protein, specifically influenza virus HA, so that it mimics a physiologically relevant trimeric viral spike (FIG. 1A). Ferritin forms a nearly spherical particle consisting of 24 subunits arranged with octahedral symmetry around a hollow interior. The symmetry of the ferritin nanoparticles includes eight three-fold axes on the surface. The aspartic acid (Asp) at residue 5 near the NH₂ terminus is readily accessible, and the distance (28 Å) between each Asp5 on the three-fold axis is almost identical to the distance between the central axes of each HA2 subunit of trimeric HA (FIG. 1A, right).

Vector Construction.

The HA-ferritin fusion proteins were constructed by joining the ectodomain of A/New Caledonia/20/1999 (1999 NC)

HA to ferritin (FIG. 1B). Specifically, the gene encoding *H. pylori* nonheme iron-containing ferritin (GenBank NP_223316) with a point mutation (N19Q) to abolish a potential N-linked glycosylation site was synthesized by PCR-based accurate synthesis (M. F. Bachmann, R. M. Zinkernagel, Neutralizing antiviral B cell responses. *Annu Rev Immunol* 15, 235-270 (1997)) using human-preferred codons. The human CD5 leader sequence and a serine-glycine-glycine (SGG) spacer were joined to the gene fragment encoding ferritin (residues 5-167) to generate a secreted protein. The plasmids encoding various influenza virus HAs, including A/South Carolina/1/1918 (1918 SC), GenBank AF117241; A/Puerto Rico/8/1934 (1934 PR8), UniProt P03452; A/Singapore/6/1986 (1986 Sing), GenBank ABO38395; A/Beijing/262/1995 (1995 Beijing), GenBank AAP34323; A/New Caledonia/20/1999 (1999 NC), GenBank AY289929; A/Solomon Islands/3/2006 (2006 SI), GenBank ABU99109; A/Brisbane/59/2007 (2007 Bris), GenBank ACA28444; A/California/04/2009 (2009 CA), GenBank ACP41105; A/Perth/16/2009 (H3 2009 Perth), GenBank ACS71642; B/Florida/04/2006 (B 2006 Florida), GenBank ACA33493 and their corresponding NAs with human preferred codons were synthesized as previously reported (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)). The gene fragments encoding HAs (residues HA1 1-HA2 174, H3 numbering) from 1999 NC HA, 2009 CA HA, 2009 Perth H3 and 2006 Florida B were amplified and joined to the ferritin gene fragment (residues 5-167) with an SGG linker to give rise to the HA-ferritin fusion gene. To produce soluble trimeric HA, the 1999 NC HA gene fragment (residues HA1 1-HA2 174, H3 numbering) was joined to a thrombin cleavage site followed by a foldon trimerization motif and a poly-histidine tag as described previously (A. S. Xiong et al., PCR-based accurate synthesis of long DNA sequences. *Nat Protoc* 1, 791-797 (2006)). Both full length and soluble forms of 1999 NC ΔStem (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)) and ΔRBS HA mutants were generated by introducing an N-linked glycosylation site at residues HA2 45 (I45N/G47T) and HA1 190 (Q192T), respectively. The soluble form of 2007 Bris ΔRBS HA mutant was generated by introducing an N-linked glycosylation site at the same site. All genes were then cloned into mammalian expression vectors for efficient expression (C. J. Wei et al., Comparative efficacy of neutralizing antibodies elicited by recombinant hemagglutinin proteins from avian H5N1 influenza virus. *J Virol* 82, 6200-6208 (2008)). Plasmids encoding the mAbs, CR6261 (D. C. Ekiert et al., Antibody recognition of a highly conserved influenza virus epitope. *Science* 324, 246-251 (2009)), CH65 (J. R. Whittle et al., Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc Natl Acad Sci USA* 108, 14216-14221 (2011)) and a single-chain variable fragment F10 (J. Sui et al., Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* 16, 265-273 (2009)) were also synthesized as described by C. J. Wei et al., (*Science* 329, 1060-1064 (2010)). Protein Biosyntheses and Purifications.

To produce ferritin nanoparticles, HA-np and trimeric HA, the expression vectors were transfected into 293F cells (Invitrogen), a human embryonic kidney cell line using 293fectin (Invitrogen) according to the manufacturer's instructions. Matched NAs were co-transfected at 20:1 HA:NA (wt:wt). The cells were grown in Freestyle 293

expression medium (Invitrogen) and the culture supernatants were collected 4 days post-transfection by centrifugation and filtered through a 0.22 µm pore filter unit (Nalgene) to remove cell debris. The supernatants were concentrated with a 30 kDa molecular weight cut-off filter unit (Pall Corp.) and then buffer exchanged to a Tris buffer (20 mM Tris, 50 mM NaCl, pH 7.5 for ferritin nanoparticles; 20 mM Tris, 500 mM NaCl, pH 7.5 for HA-np). The ferritin nanoparticles were purified by ion-exchange chromatography using a HiLoad 16/10 Q Sepharose HP column (GE Healthcare). The HA-np were purified by affinity column chromatography using *Erythrina cristagalli* agglutinin (ECA, coral tree lectin; EY Laboratories, Inc.) specific for galactose β(1,4)N-acetylglucosamine. The ferritin nanoparticles and HA-np were further purified by size exclusion chromatography with a Superose 6 PG XK 16/70 column (GE Healthcare) in PBS. The peak fraction was collected and used for further studies. The molecular weights of the ferritin nanoparticle and HA-np were calculated based on two equations generated by least squares linear regression on a semi-log plot using gel filtration low and high molecular weight standards (Bio-Rad), respectively. The yield of 1999 NC HA-np is ~4 mg liter⁻¹ and appears stable at 4° C. or frozen at -80° C. The trimeric HA proteins were purified as described by A. S. Xiong et al (*Nat Protoc* 1, 791-797 (2006)) with slight modifications. Briefly, HA proteins were first purified by affinity chromatography using Ni Sepharose HP resin (GE Healthcare), and then were separated by size exclusion chromatography with a HiLoad 16/60 Superdex 200 PG column (GE Healthcare). To remove the foldon trimerization motif and poly-histidine tag, HA proteins were digested with thrombin (EMD Chemicals, Inc.) (3 U mg ml⁻¹) overnight at 4° C. Undigested proteins were removed by passing over Ni Sepharose HP resin and the digested HAs were purified on a HiLoad 16/60 Superdex 200 PG column. All purified proteins were verified by SDS-PAGE. Protein purity and size distribution were examined by dynamic light scattering using a DynaPro system (Wyatt Technology). All human mAbs and a single-chain variable fragment were also produced in 293F cells and purified as described previously (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010); W. P. Kong et al., Protective immunity to lethal challenge of the 1918 pandemic influenza virus by vaccination. *Proc Natl Acad Sci USA* 103, 15987-15991 (2006)). MAbs against 1999 NC HA were purified from hybridoma supernatants as previously described (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)).

Iodixanol-Based Gradient Centrifugation.

Alternatively, HA-np were purified by iodixanol gradient ultracentrifugation (FIG. 10) routinely used for virus and VLP purifications (C. J. Wei et al., Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. *Sci Transl Med* 2, 24ra21 (2010)). Fractions containing HA np were confirmed by SDS-PAGE and Western blotting using a mAb against 1999 NC HA.

Electron Microscopic Analysis.

Purified ferritin nanoparticles and HA-np were subjected to transmission electron microscopic analysis. The samples were negatively stained with phosphotungstic acid (ferritin nanoparticles) or ammonium molybdate (HA-np) and images were recorded on a Tecnai T12 microscope (FEI) at 80 kV with a CCD camera (AMT Corp.).

Analysis of HA-Ferritin Np.

Among the various ferritins, *Helicobacter (H.) pylori* nonheme ferritin (K. J. Cho et al., The crystal structure of ferritin from *Helicobacter pylori* reveals unusual conformational changes for iron uptake. *J Mol Biol* 390, 83-98 (2009)) was selected as a prototype because of its highly divergent sequence compared to mammalian ferritins (FIG. 2), thus minimizing the likelihood of inducing autoimmunity after vaccination. The final purification step for recombinant HA-ferritin was size exclusion chromatography (FIG. 1C, left) and dynamic light scattering was used to confirm that both ferritin and HA-ferritin self-assembled into supramolecules with diameters of 14.61 and 37.23 nm, respectively (FIG. 1C, middle). HA-ferritin and ferritin subunits from these nanoparticles migrated at the expected respective molecular weights of 85 and 17 kDa by SDS-PAGE compared to 68 kDa for purified HA (FIG. 1C, right). While the morphology of the ferritin nanoparticles was smooth, as visualized by transmission electron microscopy (TEM), HA-ferritin formed np that exhibited clearly visible spikes around the spherical core (FIG. 1D, Ferritin np vs. HA-np). Remarkably, the placement of these spikes clearly illustrated the octahedral symmetry of the HA-np design. Octahedral two-, three- and four-fold axes were distinctly observed in the TEM image (FIG. 1E, right). These data demonstrated the formation of HA spikes on self-assembling HA-ferritin nanoparticles. More importantly, this design enabled HAs from different subtypes or influenza B viruses to be readily joined to a ferritin core without substantial modification.

Example 2

Antigenicity and Immunogenicity of HA-Np in Mice

To verify the antigenicity of the HA spikes on the np, HA-ferritin np were analyzed for their ability to react with anti-HA head ab and a conformation-dependent monoclonal ab (mAb), CR6261, that recognizes a highly conserved structure in the trimeric HA stem and neutralizes diverse influenza A group 1 viruses D. C. Ekiert et al., Antibody recognition of a highly conserved influenza virus epitope. *Science* 324, 246-251 (2009)), using ELISA and a virus neutralization assay.

Analysis by ELISA.

Purified trimeric HA, HA-np, and TIV (2 µg of H1 HA ml⁻¹), ferritin nanoparticles (0.68 µg ml⁻¹ for FIG. 3 or 2 µg ml⁻¹ for the rest), mouse liver ferritin (2 µg ml⁻¹, Alpha Diagnostic International, Inc.), ΔStem and ΔRBS HA trimer (2 µg ml⁻¹) were coated (100 µl/well) onto MaxiSorp™ plates (Nunc) and the wells were probed with the anti-HA mAbs, anti-mouse liver ferritin IgG (Alpha Diagnostic International, Inc.) or immune sera followed by peroxidase-conjugated secondary antibodies (anti-mouse IgG and anti-human IgG, SouthernBiotech; anti-ferret IgG, Rockland Immunochemicals, Inc.). The wells were developed using a SureBlue chromogen (KPL) and the reaction was stopped by adding 0.5 M sulfuric acid. For the ELISA-based competition assay, HA trimer (2 µg ml⁻¹) was coated onto the plates. Plates were incubated with an anti-stem mAb, CR6261 (8 µg ml⁻¹) or an isotype control Ab, VRC01 (8 µg ml⁻¹) (Z. Y. Yang et al., Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science* 317, 825-828 (2007); X. Wu et al., Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 329, 856-861 (2010)) before adding serially diluted pre-absorbed ferret immune sera. The

wells were probed with anti-ferret IgG and developed as described above. Absorbance at 450 nm was measured by SpectraMax M2e (Molecular Devices). The endpoint titers were determined by calculating the intersection of the observed binding curve and the absorbance threshold (four times background).

Neutralization Assays.

HA/NA-pseudotyped lentiviral vectors encoding luciferase were used. Immune sera used for the assay were pretreated with RDE as described above. Pre-titrated pseudotyped viruses (Gag p24 \approx 6.25 ng ml $^{-1}$) were incubated with serially diluted sera for 20 minutes at room temperature and added to 293A cells (10,000 cells/well in a 96-well plate; 50 μ l/well; in triplicate). Plates were then washed and replaced with fresh media 2 hours later, and luciferase activity was measured after 24 hours. For the protein competition assay, neutralizing activity of the mAbs F10, CR6261 or immune sera was measured in the presence of competitor proteins, trimeric HA (WT, Δ Stem or Δ RBS), HA-np, ferritin nanoparticles or irrelevant protein (HIV-1 gp120) at final concentration of 20 and 25 μ g ml $^{-1}$ for mAbs and immune sera, respectively. The HA-np was able to bind to anti-head or anti-stem mAbs with affinities similar to trimeric HA or trivalent inactivated vaccine (TIV) containing the same 1999 NC HA at equimolar concentrations of HA, in contrast to a ferritin nanoparticle control (FIG. 3A). Analogous to trimeric HA, the HA-np also blocked neutralization by CR6261 and another stem-directed mAb, F10 (4 J. Sui et al., Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* 16, 265-273 (2009)) (FIG. 3B). These results indicated that HA molecules on the HA-np antigenically resembled the physiological trimeric viral spike.

Example 3

Immunogenicity of HA-Ferritin Np In Vivo

This Example demonstrates the ability of HA-ferritin np of the present invention to elicit neutralizing antibodies.

To assess the immunogenicity of the HA-ferritin np in vivo, mice were immunized twice with HA-np or TIV's from the 2006-2007 season, with HAs from A/New Caledonia/20/1999 (H1N1), A/Wisconsin/67/2005 (H3N2) and B/Malaysia/2504/04 (type B), or from the 2011-2012 season, with HAs from A/California/07/09-like (H1N1), A/Perth/16/09 (H3N2) and B/Brisbane/60/08 (type B). Briefly, female BALB/c mice (6-8 weeks old; Charles River Laboratories) were immunized (5 mice/group) intramuscularly with 5 or 0.5 μ g (1.67 or 0.17 μ g of H1 HA) of TIV, 2.24 or 0.22 μ g (1.67 or 0.17 μ g of HA) of HA-np or 0.57 μ g of ferritin nanoparticles (equimolar to 2.24 μ g of HA-np) in 100 μ l of PBS or in 100 μ l of 50% (v/v) mixture of Ribi adjuvant (Sigma) in PBS at weeks 0 and 3. A group of BALB/c mice (n=4) was immunized with 20 μ g of trimeric HA (thrombin cleaved) in 100 μ l of 50% (v/v) mixture of Ribi adjuvant in PBS at weeks 0 and 4. For the experiment using trivalent HA-np, mice were immunized (n=5) with 6.72 μ g (1.67 μ g of each HA component) of trivalent HA-np in 100 μ l of 50% (v/v) mixture of Ribi adjuvant in PBS at weeks 0 and 3. Blood samples were collected prior to the first dose, and at 2 weeks after each immunization.

The resulting antibody titers were determined as described in Example 2. The HA-np induced significantly higher HAI titers than TIV (FIG. 4A, left; p<0.0001), and a similar effect was observed in the neutralization assay and

ELISA (FIG. 4A, middle and right; p<0.0001). For example, neutralization titers elicited by HA-np as assessed by the concentration of ab needed to inhibit viral entry by 90% (IC_{90}) were ~34 times higher than TIV (FIG. 4A, middle). Because higher titers were observed in groups with the adjuvant Ribi, further comparisons were performed with this adjuvant. Neutralization against a panel of H1N1 strains revealed not only increased potency but also enhanced breadth stimulated by HA-np compared with TIV or trimeric HA (FIG. 4B). Neutralization against two highly divergent H1N1 viruses, A/Puerto Rico/8/1934 (1934 PR8) and A/Singapore/6/1986 (1986 Sing) were only observed in mice immunized with the HA-np, and the titer against the contemporary virus A/Brisbane/59/2007 (2007 Bris) was more than one log higher in mice immunized with HA-np than with TIV (FIG. 4B).

To assess whether the preexisting immune responses to ferritin nanoparticles or to other HA subtypes would attenuate the immunogenicity of the subsequent immunization of HA-np, mice were pre-immunized with either H3 (A/Perth/16/09, 2009 Perth) HA-np or empty ferritin nanoparticles to elicit anti-H3 HA and/or anti-*H. pylori* ferritin immune responses (FIG. 5A). These animals were then immunized with H1 (1999 NC) HA-np. Comparable HAI, IC_{90} neutralization and ELISA titers against 1999 NC HA were observed in naïve animals as well as in groups pre-immunized with H3 HA-np or empty ferritin nanoparticles (FIG. 5B). These results indicated that preexisting anti-*H. pylori* ferritin immunity did not diminish the HA-specific ab response.

Example 4

Lack of Autoreactivity of *H. pylori* Ferritin Nanoparticles

This Example demonstrates analyzes the ability of HA-ferritin np of the present invention to elicit an auto-immune response against autologous ferritin in mice.

Although the overall structural architecture and physiological functions of ferritin are conserved across organisms, murine ferritin has only 27% amino acid sequence identity to *H. pylori* ferritin. This homology nonetheless raised the possibility that immunization with *H. pylori* ferritin in mice might abrogate immune tolerance and induce autoimmunity. To address this concern, CD4, CD8 T-cell and ab responses against both murine and *H. pylori* ferritins were analyzed by intracellular cytokine staining (ICS) and ELISA in mice immunized with HA-np. ELISAs were performed according to the procedure in Example 2. For intracellular cytokine analysis, CD4 $^{+}$ and CD8 $^{+}$ T-cell responses were evaluated for interferon- γ (IFN- γ), tumor necrosis factor α (TNF α), and interleukin-2 (IL-2) as described by T. Zhou et al. (*Science* 329, 811-817 (2010)). Individual peptide pools (15-mer overlapping by 11 residues, 2.5 μ g ml $^{-1}$ for each peptide) covering *H. pylori* ferritin or mouse ferritin light and heavy chains were used to stimulate cells. After stimulation, cells were fixed, permeabilized and stained using anti-mouse CD3, CD4, CD8, IFN- γ , TNF α and IL-2 mAbs (BD Pharmingen) together with aqua blue dye for live/dead stain (Invitrogen). The data were collected by LSR II Flow Cytometer (BD Biosciences) and IFN- γ , TNF α - and IL-2-positive cells in the CD4 $^{+}$ and CD8 $^{+}$ cell populations were analyzed with FlowJo software (Tree Star).

Although an increase in the ICS staining of CD4 $^{+}$ T cells stimulated with *H. pylori* ferritin peptides (FIG. 4C, top left) was observed, no increases in the CD4 $^{+}$ and CD8 $^{+}$ ICS

59

responses were seen with murine ferritin peptide stimulation (FIG. 4C, bottom left and middle). In addition, while high titers ($>10^6$) of anti-*H. pylori* ferritin abs were detected in ferritin nanoparticle- and HA-np-immune sera, abs to mouse ferritin were undetectable (FIG. 4C, right). These results demonstrate that HA-ferritin np of the present invention do not elicit autoreactivity to autologous ferritin in mice.

Example 5

Generation of Trivalent HA-Np and Immunogenicity in Mice

The Example analyzes whether multivalent HA-np were similar in immunogenicity to monovalent np.

HA-np expressing HAs from H1 (A/California/04/09, 2009 CA), H3 (2009 Perth) or influenza B (B/Florida/04/06, 2006 FL) were generated. The 2009 CA (H1)-, 2009 Perth (H3)- and 2006 FL (type B)-HA-np self-assembled and displayed the same morphology observed for 1999 NC HA-np (FIG. 6A). Trivalent HA-np were generated by combining three monovalent HA-np, and their immunogenicity was compared to a seasonal TIV containing the same H1 and H3 strains and a mismatched type B (B/Brisbane/60/08). HAI titers against homologous H1N1 and H3N2 viruses were significantly increased in animals immunized with trivalent HA-np relative to TIV-immunized animals (FIG. 6B; $p=0.0125$ and 0.0036 , respectively). When compared to animals immunized with the corresponding monovalent HA-np, HAI titers against 2009 CA (H1) and 2009 Perth (H3) induced by trivalent HA-np were comparable (FIG. 6B). These results demonstrate that no substantial antigenic competition between H1 and H3 HA-np was observed with a trivalent HA-np vaccine.

Example 6

Cross-Protective Immunity Elicited by HA-Np in Ferrets

This Example demonstrates that vaccination of ferrets with 1999 NC HA-np elicits a protective immunity similar to that observed in human disease.

Male Fitch ferrets (6 months old; Triple F Farms), seronegative for exposure to H1N1, H3N2 and type B influenza viruses, were housed and cared for at BIOQUAL, Inc. (Rockville, Md.). Prior to study start, a temperature transponder (Biomedic Data Systems, Inc.) was implanted into the neck of each ferret. Ferrets were immunized (6 ferrets/group) intramuscularly with 500 μ l of PBS, 7.5 μ g (2.5 μ g of H1 HA) of TIV or 3.35 μ g (2.5 μ g of HA) of HA-np in 500 μ l of 50% (v/v) mixture of Ribi adjuvant in PBS at weeks 0 and 4. Blood was collected 3 and 2 weeks after the first and the second immunization, respectively.

Three weeks after the first immunization, all ferrets receiving HA-np generated protective HAI titers against homologous H1 1999 NC virus ($>1:40$), while only 50% (3/6) of TIV-immunized ferrets induced HAI titers greater than 1:40 (FIG. 7A, left; $p=0.0056$). The same trend was also observed for both neutralization and anti-HA ab titers (FIG. 7A, middle and right; $p=0.0047$ and $p=0.0045$, respectively), documenting the superior potency of HA-np in a second species. After boosting, the HAI and IC₉₀ neutralization titers of the HA-np-immune sera were ~10-fold higher than those of TIV-immunized ferrets (FIG. 7A, left and middle; 457 ± 185 vs. 5760 ± 1541 , $p=0.0066$, and 598 ± 229 vs. 5515 ± 1074 , $p=0.0012$, respectively). A similar enhancement

60

in HA-np vs. TIV immunization was also observed by ELISA titers (FIG. 7A, right; $p=0.0038$). Remarkably, a single immunization with HA-np induced immune responses comparable to two immunizations with TIV (FIG. 7A).

To determine whether HA-np could confer protection against an unmatched H1N1 virus, five weeks after the last immunization ferrets immunized with 1999 NC HA-np or TIV containing the same H1 HA were challenged with $10^{6.5}$

¹⁰ EID₅₀ of 2007 Bris virus. (1999 NC and 2007 Bris viruses are 8 years apart and their antigenic characteristics are sufficiently different to require the production of two different vaccines to confer protection in humans.) The virus was expanded in embryonated chicken eggs from a seed stock obtained from CDC (Atlanta, Ga.) and has a titer of $10^{6.5}$ ¹⁵ EID₅₀ ml⁻¹. The virus stock was inoculated intranasally into ferrets, which had been anesthetized with ketamine/xylazine, in a volume of 500 μ l per nostril. The ferrets were observed for clinical signs twice daily and weight and ²⁰ temperature measurements recorded daily by technicians blind to the treatment groups. Nasal washes were obtained on days 1, 3 and 5 and infectious viral titers were determined by TCID₅₀ assay using MDCK cells as described previously (C. J. Wei et al., Induction of broadly neutralizing H1N1 ²⁵ influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)).

Ferrets immunized with HA-np showed a significant reduction in viral shedding beginning 1 day after challenge compared to the sham control group (FIG. 7B, left; ³⁰ $p=0.0259$). At the same time point, no reduction in viral shedding was seen in the TIV-immunized group. Four of six animals immunized with HA-np had no detectable viral load after 3 days and by day 5, all animals in this group cleared the virus, while all animals in the sham control group still ³⁵ had detectable virus (FIG. 7B). In addition, HA-np-immunized ferrets suffered less body weight loss compared to the TIV-immunized and sham control groups (FIG. 7B, right). These results demonstrate faster virus clearance in ferrets immunized with HA-np than with TIV and further demonstrate that HA-np effectively induced cross-protective ⁴⁰ immunity in vaccinated ferrets.

Example 7

Induction of Two Types of Neutralizing Abs (nAbs) in Ferrets

This Example demonstrates the breadth and specificity of nAbs in ferret immune sera.

⁴⁵ IC₅₀ neutralization titers against 1986 Sing, A/Beijing/262/1995 (1995 Beijing), A/Solomon Islands/3/2006 (2006 SI) and 2007 Bris were significantly higher in animals immunized with HA-np compared to immunization with TIV (FIG. 8A, left). This enhanced breadth was due not only ⁵⁰ to a quantitative increase in overall ab titer (~9-fold against matched virus) but also reflected a qualitative difference in the types of abs elicited (>40-fold enhancement against an unmatched strain). To determine whether the cross-reactivity induced by HA-np was due to nAbs to the conserved HA stem epitope, ferret immune sera were pre-absorbed with cells expressing a stem mutant (Δ Stem) HA to remove non-stem directed antibodies. Briefly, ferret immune sera taken 2 weeks after the second immunization were subjected to the assay. The plasmids encoding for Δ Stem and ARBS ⁵⁵ HAs were transfected into 293F cells. Three days after transfection, the cells were analyzed by flow cytometry to confirm expression of HA on the cell surface and used for ⁶⁰

61

serum absorption. One ml of the immune sera diluted at 1:100 and 1:1,000 was incubated with 100 μ l of pre-washed Δ Stem and Δ RBS HA-expressing 293F cell pellets, respectively. After incubating for 1 hour at 4°C., supernatants were harvested by centrifugation and binding to WT and mutant HAs was examined by ELISA previously described (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)). The Δ Stem HA-pre-absorbed sera were also used for competition ELISA.

Stem-specific abs were detected in HA-np-immunized ferrets (6/6) in greater frequency and magnitude than TIV-immune ferrets (2/6) (FIG. 8B, left; p=0.0056). Moreover, binding of these pre-absorbed sera to HA was inhibited by CR6261 mAb (FIG. 8B, right; p=0.0019), further documenting the specificity of HA-np immune sera to the stem epitope. The HAI titers against heterologous 2007 Bris virus were also significantly higher in ferrets immunized with HA-np (6/6, 1:80-1:640) than with TIV (3/6, 1:40-1:80) (FIG. 8A, right; p=0.0054). Interestingly, in contrast to a previous study in which DNA prime/TIV boost was used to elicit anti-stem broadly neutralizing abs (bnAbs) (C. J. Wei et al., Induction of broadly neutralizing H1N1 influenza antibodies by vaccination. *Science* 329, 1060-1064 (2010)), sera from animals immunized with HA-np showed HAI ab titers against a highly divergent 1934 PR8 strain, with titers \geq 1:40 in all ferrets. However, no HAI titers against 1934 PR8 were detected in TIV-immunized ferrets (FIG. 8A, right). These data suggested that the HA-np vaccine might elicit another class of nAb directed towards the conserved RBS in the HA head.

To determine whether HA-np elicited abs against RBS, an RBS mutant HA (Δ RBS) was generated by introducing a glycosylation site in the sialic acid binding pocket at residue 190 (FIG. 9) (D. Lingwood et al., Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature*, in press). Ferret immune sera were absorbed with Δ RBS HA-expressing cells to remove abs to HA outside of this region and tested for binding against WT or Δ RBS HA. RBS-directed abs were detected with titers of >1:2,000 in all HA-np-immunized ferrets, but only 1 out of 6 ferrets that received TIV (FIG. 8B, middle).

To define the relative contributions of these stem and RBS abs to the breadth of neutralization, neutralization assays were performed in the presence of competitor proteins: WT, Δ Stem or Δ RBS HA. In the presence of excess Δ Stem HA, only stem-directed abs can neutralize viruses; similarly, Δ RBS HA interferes with all antibodies in the serum except those proximal to the RBS. The relative contribution of stem- and RBS-directed neutralization was measured as activity remaining in the presence of the respective competitor HA. For example, with 2007 Bris, Δ RBS HA only partially inhibited neutralization, while either WT or Δ Stem HA almost completely abolished the neutralization activity of the sera; hence, the neutralization against 2007 Bris was due almost entirely to RBS-directed abs (FIG. 8C). Four H1N1 strains were tested in this assay. The pattern of neutralization inhibition varied by strain. Neutralization of 1999 NC or 2007 Bris was mediated predominantly by RBS-directed abs. However, neutralization of 1986 Sing was due mainly to stem-directed abs. Interestingly, the neutralization of 1995 Beijing was more complex. Both stem- and RBS-directed abs contributed to neutralization of this virus (FIG. 8C).

These results demonstrate that HA-np induce both known types of bnAbs—stem-directed and RBS-directed. Together, these abs contribute to the breadth and potency of the

62

immune sera elicited by HA-np. The synergy between them explains mechanistically the observed superior efficacy of the HA-np vaccine and decreases the likelihood of viral escape mutations from either antibody alone.

Taken together the above-disclosed Examples demonstrate that a ferritin-based nanoparticle is able to present trimeric HA in its native fold, rigidly and symmetrically, with sufficient spacing to ensure optimal access to potential bnAbs directed to the stem. They also demonstrate that the nanoparticles have enhanced immunogenicity and an expanded neutralization breadth to both stem and RBD antibodies.

Example 8

Immunization of Mice and Ferrets Using a Tetraivalent Vaccine

This Example demonstrates the ability of a multivalent vaccine to elicit an immune response against several strains and sub-types of influenza virus.

The ability of a pan-group 1 vaccine to stimulate neutralizing antibodies against a variety of influenza viruses was tested in mice and ferrets using a protocol similar to that described in Example 1, and outlined in FIG. 11. Briefly, a pan-group 1 HA-ferritin np vaccine was produced by combining four different monovalent HA-ferritin np vaccines. Specifically, HA-ferritin np, each expressing either H1 A/NC/20/1999, H1 A/CA/04/2009, H2 A/Singapore/1/1957 or H5 A/Indonesia/05/2005, were combined to produce a single vaccine containing all four HA proteins. Mice were immunized twice in a four week interval using 6.8 ug total of the pan-group 1 vaccine (1.7 ug of each HA-ferritin np) in Ribi. Ferrets were immunized twice in a four week interval using 10 ug total of the pan-group 1 vaccine (2.5 ug of each HA-ferritin np) in Ribi. Blood was obtained from the immunized animals and the titer of neutralizing antibodies against various influenza viruses measured. The results of this analysis are shown in FIGS. 12-14. Immunized ferrets were also challenged with either influenza A/Brisbane/59/2007 Brisbane (H1N1) (2207 Bris) (FIG. 15) or influenza A/Mexico/2009 (H1N1) (2009 Mex) (FIG. 16) and the resulting virus titers measured on day 3 and 5 post-challenge.

Example 9

Design and Construction HA-Ferritin Stem-Region Fusion Proteins

This Example demonstrates the construction of HA-ferritin proteins and nanoparticles that present the stem region of the influenza HA protein.

As illustrated in FIG. 17, the stem region of the influenza HA protein is highly conserved among different influenza strains, and possesses a site of vulnerability for Group 1 viruses. Thus, a vaccine that elicits neutralizing antibodies against the stem region of the influenza HA protein should be broadly neutralizing. A nanoparticle displaying the stem region of the influenza stem region was constructed as a vaccine.

Design of an HA-Stabilized Stem Fusion Protein.

An HA-stabilized stem fusion protein (HA SS) was constructed as follows: residues 43-313 of the head domain of HA1 were replaced with a Gly-Trp-Gly linker. The membrane distal end of HA2 (residues 59 to 93) was replaced by an HIV-1 Bal gp41 HR2 helix followed by a six residue

63

glycine-rich linker (Asn-Gly-Thr-Gly-Gly-Ser-Gly) and the gp41 HR1 helix. The HR1 helix of gp41 was added in frame with helix C of HA2 so as to generate a long central chimeric helix. The resulting six helix bundle sitting atop the modified hemagglutinin stem provides stability to the SS trimer in lieu of the missing head residues. A schematic of the resulting protein is shown in FIG. 18A, while a ribbon diagram is shown in FIG. 18B. A second trimerization domain consisting of a 28 residue T4 foldon domain was joined to the membrane proximal C-terminus of HA2. The HA SS-ferritin nanoparticle (HA SS-*np*) protein was generated by joining residue 174 (H3 numbering) of HA SS to *H. pylori* ferritin (residues 5-167) with a Ser-Gly-Gly linker.

In constructing HA-SS fusion proteins, genes encoding wild-type HA proteins (A/Puerto Rico/8/1934 (H1 1934 PR8), A/Singapore/6/1986 (H1 1986 Sing), A/New Caledonia/20/1999 (H1 1999 NC), A/Brisbane/59/2007 (H1 2007 Bris), A/Vietnam/1203/2004 (H5 2004 VN), A/Canada/720/05 (H2 2005 CAN), A/Hong Kong/I/1968 (H3 1968 HK), A/Hong Kong/1073/1999 (H9 1999 HK) and their corresponding NAs, H1 NC 99 SS, RSC3 HIV gp120 control protein, and all Abs (CR6261, F16v3, and VRCo1) were synthesized with human preferred codons as previously described (Wei et al. Science 2010, 329(5995):1060-4). *Helicobacter pylori* nonheme iron-containing ferritin (GenBank NP_223316) with a point mutation (N19Q) to abolish a potential N-linked glycosylation site was synthesized by PCR-based accurate synthesis (Xiong et al. Nat Protoc 2006, 1(2):791-797) using human-preferred codons. Coding sequences for the human CD5 leader sequence and a serine-glycine-glycine (SGG) spacer were joined to the gene fragment encoding ferritin (residues 5-167) to generate a secreted protein. HA and HA SS-*np* fusion proteins were generated by overlap PCR by joining the HA ectodomains at residue HA2 174 (H3 numbering) to *H. pylori* ferritin (residues 5-167) with a Ser-Gly-Gly linker. Stem mutant probes Δstem (glycosylation insertion into the CR6261 binding epitope at position 45 in HA2; H3 numbering) which prevent binding at the conserved H1 stem epitope were generated using site directed mutagenesis. Genes encoding these proteins were cloned into a CMVR plasmid backbone for efficient mammalian cell expression.

Protein Expression and Purification

Plasmids encoding soluble proteins were transfected (HA ectodomain genes were cotransfected with the corresponding NA encoding plasmids) into the human embryonic kidney cell line 293F and isolated from expression supernatants 72-96 hrs post-transfection. All HA and HA SS trimeric proteins were purified first by metal chelation affinity chromatography and then by size exclusion chromatography as previously described (Wei et al. J Virol. 2008, 82(13):6200-8). IgG Abs were purified using a Protein G affinity column (GE Healthcare). The HA- and HA SS-*np* were purified by affinity column chromatography using *Erythrina crista-galli* agglutinin (ECA, coral tree lectin; EY Laboratories, Inc.) specific for galactose β(1,4)N-acetylglucosamine and *Galanthus nivalis* agglutinin (GNA, snowdrop lectin; EY Laboratories, Inc.) specific for α(1,3) and α(1,6) linked high mannose structures, respectively. HA- and HA SS-*np* were further purified by size exclusion chromatography with a Superose 6 PG XK 16/70 column (GE Healthcare) in PBS (FIG. 19).

HA SS-Ferritin Characterization.

HA SS-ferritin *np* were visualized by electron microscopy. Briefly, purified HA SS-*np* were negatively stained with phosphotungstic acid and ammonium molybdate, respectively, and images were recorded on a Tecnai T12

64

microscope (FEI) at 80 kV with a CCD camera (AMT Corp.). The results of this analysis are shown in FIG. 20. IN addition, the ability of purified HA SS and HA SS-*np* to bind to monoclonal Abs CR6261 and F16v3 (1.7×10^{-4} to 10 µg/mL) was characterized by ELISA. HA and HIV gp120 proteins served as controls. Ab binding was detected by peroxidase-conjugated goat anti-human IgG. The results of this analysis, which are shown in FIG. 21, demonstrate that HASS-ferritin is antigenically similar to HA protein.

Example 10

Immune Response to HA SS-Ferritin Nanoparticles

This Example demonstrates the immune response generated in animals following immunization with HA SS-ferritin *np*.

BALB/c mice were immunized twice intramuscularly with protein (2 or 10 µg each) formulated with Ribi adjuvant system (Sigma) at a 3 week interval. Mice received either homologous (HA SS-*np* prime and boost) or heterologous (HA-*np* prime and HA SS-*np* boost) immunizations. Ferrets were immunized three times intramuscularly with HA SS-*np* (10 µg each) formulated with Ribi adjuvant system (Sigma) at weeks 0, 4 and 14. Serum was collected from animals 2 weeks after each immunization and 1 week prior to the first immunization and heat inactivated (30 min at 56° C.).

Pre- and post-immune sera from immunized mice and ferrets were assayed for binding to HA and HA SS by ELISA. Briefly, sera were serially diluted (diluted 50 to 2.3×10^6) and assayed for reactivity to soluble trimeric HA and HA SS proteins, as well as control proteins (200 ng/well with molar equivalents plated according to HA SS). Binding was detected by peroxidase conjugated anti-mouse or anti-ferret IgG, respectively. Endpoint dilutions were determined from nonlinear fit dose-response curves using a detection limit of 2× background absorbance. The result from this analysis are shown in FIG. 22 and demonstrate that stem specific cross-reactive antibodies which recognize the conserved stem-epitope are elicited by HA SS-*np* vaccination.

Sera were also analyzed for neutralization of pseudotyped recombinant lentiviruses expressing wild-type HA with the corresponding NA with a luciferase reporter gene as previously described (Wei et al. Science 2010, 329(5995):1060-4) following pretreatment with receptor-destroying enzyme (RDE II; Denka Seiken Co., Ltd.). Pseudotype neutralization competition of ferret serum was performed by incubating serially diluted serum in the presence of either H1 1999 NC SS, H1 1999 NC SS Δstem probe or gp120 control (10 µg/mL) for 1 hr (RT) before addition to pseudotyped recombinant lentiviruses and assaying for neutralization. The results from this analysis are shown in FIG. 23 and demonstrate that vaccination with HA SS-*np* elicits neutralizing antibodies against various group-1 strains.

Example 11

Immune Response to HA SS-Ferritin Heterologous Immunization Boost

This example demonstrates that HA SS-*np* can be utilized to boost antibodies directed to the conserved stem epitope.

BALB/c mice were immunized twice intramuscularly with heterologous ferritin proteins (HA-*np* prime and HA SS-*np* boost; 2 µg each) formulated with Ribi adjuvant system (Sigma) at a 3 week interval. Serum was collected

65

from animals 2 weeks after each immunization and 1 week prior to the first immunization and heat inactivated (30 min at 56° C.).

Pre- and post-immune sera from immunized mice were assayed for binding to HA and HA SS by ELISA. Briefly, sera were serially diluted (diluted 50 to 2.3×10^6) and assayed for reactivity to soluble trimeric HA and HA SS proteins, as well as control proteins (200 ng/well with molar equivalents plated according to HA SS). Binding was detected by peroxidase conjugated anti-mouse or anti-ferret IgG, respectively. Endpoint dilutions were determined from nonlinear fit dose-response curves using a detection limit of $2\times$ background absorbance. The results from this analysis are shown in FIG. 22 and demonstrate that cross-reactive stem-epitope specific antibodies are being elicited.

Sera were also analyzed for neutralization of pseudotyped recombinant lentiviruses expressing wild-type HA with the

66

corresponding NA with a luciferase reporter gene as previously described (Wei et al. *Science* 2010, 329(5995):1060-4) following pretreatment with receptor-destroying enzyme (RDE II; Denka Seiken Co., Ltd.). The results from this analysis are shown in FIG. 24 and demonstrate that mice which have preexisting stem antibodies titers can be boosted with HA SS-np.

While the present invention has been described with reference to the specific embodiments thereof, it should be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the true spirit and scope of the invention. In addition, many modifications may be made to adapt a particular situation, material, composition of matter, process, process step or steps, to the objective, spirit and scope of the present invention. All such modifications are intended to be within the scope of the claims.

SEQUENCE LISTING

-continued

Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His
 115 120 125

Glu Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile
 130 135 140

Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly
 145 150 155 160

Ile Ala Lys Ser Arg Lys Ser
 165

<210> SEQ ID NO 3

<211> LENGTH: 504

<212> TYPE: DNA

<213> ORGANISM: Helicobacter pylori

<400> SEQUENCE: 3

ttaagatttc ctgcttttag cgatccctt gacatactga tcggctaaat acaaggccatg	60
gttttcatta ccaatcaact caattttac caaaaatatcc ttgaaaagca cttcttc	120
atgcgtgttca gccacatacc attgcaagaa attgaaagtgc gcatgatctt tgcttttat	180
ggcgtgtatct acgtatattgt taatagactc gctgatgtgt tgctcatgtt cataggctt	240
ttggaaaatt tgagtcaaactt cttcaaaactt atgctcaggc gcgctgtatgc tggcaattt	300
cacaggcaca ttgtttcat tcaagaagat aataagcttt ttagcatgct cgtattttc	360
agccgcatgg tcaaacaaga aaagccccgc gccatctaag ctatggtat agcaccatga	420
actcatgctc atatacaagt tggaaagagtt catttcctta ttcaacttgtt cgtttagcaa	480
cttaatgatg tcttttgcata acat	504

<210> SEQ ID NO 4

<211> LENGTH: 489

<212> TYPE: DNA

<213> ORGANISM: Helicobacter pylori

<400> SEQUENCE: 4

gacatcatca agctgctgaa cgagcagggtg aacaaggaga tgcagagcag caacctgtac	60
atgagcatga gcagctggtg ctacacccac agcctggacg ggcggccct gttcctgttc	120
gaccacgccc ccgaggagta cgagcacgccc aagaagctga tcatttctt gaacgagaac	180
aacgtgccc tgcagctgac cagcatcgc gccccggc acaagttcgaa gggctgacc	240
cagatcttcc agaaggccta cgagcacgag cagcacatca gcgagagcat caacaacatc	300
gtggaccacg ccatcaagag caaggaccac gccaccttca acttcctgca gtggatcg	360
gecgagcagc acgaggagga ggtgctgttc aaggacatcc tggacaagat cgagctgatc	420
ggcaacgaga accacggcct gtacctggcc gaccagtagc tgaagggcat cgccaagagc	480
aggaagagc	489

<210> SEQ ID NO 5

<211> LENGTH: 163

<212> TYPE: PRT

<213> ORGANISM: Helicobacter pylori

<400> SEQUENCE: 5

Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln Ser
 1 5 10 15

Ser Asn Leu Tyr Met Ser Met Ser Trp Cys Tyr Thr His Ser Leu
 20 25 30

Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu Tyr Glu

US 9,441,019 B2

69**70**

-continued

35 40 45

His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro Val
 50 55 60

Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu Thr
 65 70 75 80

Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser Glu Ser
 85 90 95

Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala Thr
 100 105 110

Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu Val
 115 120 125

Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn
 130 135 140

His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser
 145 150 155 160

Arg Lys Ser

<210> SEQ ID NO 6

<211> LENGTH: 489

<212> TYPE: DNA

<213> ORGANISM: Helicobacter pylori

<400> SEQUENCE: 6

gcttttctg ctcttggcga tgcccttcac gtactggtcg gcccgtaca ggccgtgggt 60
 ctcgttgcgc atcagctcga tcttgtccag gatgtccttg aacagcacct cctccctcg 120
 ctgcteggcc acgttaccact gcaggaagt gaagggtggcg tggtccttgc tcttgc 180
 gtggtccacg atgttgttga tgctctcgat gatgtgtcg tcgtgtcgat aggccctcg 240
 gaagatctgg gtcaggccct cgaacctgtg ctggggggcg ctgatgtgg tcagtcac 300
 gggcaegttt ttctcggtca ggaagatgtat cagtttcttgc gctgtgtcgat actccctcg 360
 ggccgtgtcg aacaggaaca ggccggcgcc gtccaggctg tgggtgttagc accagctgt 420
 catgctcatg tacagggttgc tgctctgcat ctccctgttc acctgtcgat tcagcagctt 480
 gatgatgtc 489

<210> SEQ ID NO 7

<211> LENGTH: 1695

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 7

atgaaggcca aactgctgggt gctgctgtgt acctttaccc ccacccatgc cgacacaatc 60
 tgtatcggtt accacgccaa caatagcacc gacaccgtgg atacagtgt ggagaagaac 120
 gtgaccgtga cccactctgt gaacctgtcg gaggacagcc acaatggcaa gctgtgtctg 180
 ctgaaaggca ttgcccccttgc agctgtggc aattgttctg tggccggatg gattctggc 240
 aaccccgagt gtgagctgt gatttctaag gagagctgg gctacatcgat ggagaccccc 300
 aatccctgaga atggcaccctg ctaccctggc tacttcgtcc attacgagga gctgtgtcg 360
 cagctgtcta gcgtgtccag ctccgagaga ttccgagatct tccccaaggaa gtccagctgg 420
 cctaattcaca cagtgcacagg cgtgtctgcc agctgttagcc acaacggcaa aagcagcttc 480
 taccggaaacc tgctgtggct gacaggcaag aatggcctgt accccaaaccc gagcaagac 540
 tacgtgaaca acaaggaaaa ggaagtgtcg gtgtgtggg gagtgcacca ccctccaaac 600

-continued

atcgaaatc agcgccct gtaccacaca gagaacgcct atgtgagcgt ggtgtccagc	660
cactacagca gaagattcac ccccgagatc gccaagagac ccaaagttag agaccaggag	720
ggccggatca attactactg gacctgctg gagcctggcg ataccatcat ctgcaggcc	780
aacggcaatc tgatcgcccc ttggtatgcc tttgcctgta gcagaggcctt tgccagccgc	840
atcatcacaa gcaacgcccc catggatgag tgtgatgcca agtgccagac acctcaggc	900
gccccata gcagectgcc cttccagaat gtgcacccctg tgaccatcg cgagtgc	960
aagtatgtga gaagcgccaa gctgagaatg gtgaccggcc tgagaaacat ccctagcatc	1020
cagagcagag gactgtttgg agccatcgcc ggattcatcg agggaggatg gacaggcatg	1080
gtggatggct ggtacggcta ccaccaccag aatgagcagg gctctggata tgccgccc	1140
cagaagtcta cccagaacgc catcaacggc atcaccacaa aggtgaacag cgtgatcg	1200
aagatgaaca cccagttac cgctgtggcc aaggagttca acaagctgga gcggaggatg	1260
gagaacctga acaagaaggt ggacgacggc tttctggaca tctggaccta caatgccaa	1320
ctccctggtcc tcctcgagaa tgagaggacc ctggacttcc acgacacgaa cgtgaagaac	1380
ctgtatgaga aggtgaagag ccagctgaag aacaacgcca aggagatcg caacggctgc	1440
ttcgagttct accacaagtta aacaacgag tigtatggaga gcgtgaagaa cggcacctac	1500
gactacccta agtacagcga ggagagcaag ctgaaccggg agaagatcga tggcgtgaag	1560
ctggagagca tgggcgtgta tcagatccctg gccatctaca gcacagtggc ctcttctcg	1620
gtgctgctgg tgtctctggg cgccatctcc tttggatgt gctccaacgg cagcctgcag	1680
tgcaggatct gtatc	1695

<210> SEQ_ID NO 8

<211> LENGTH: 565

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 8

Met	Lys	Ala	Lys	Leu	Leu	Val	Leu	Leu	Cys	Thr	Phe	Thr	Ala	Thr	Tyr
1									10						15

Ala	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
			20						25						30

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
			35												45

Leu	Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile
			50												60

Ala	Pro	Leu	Gln	Leu	Gly	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly
			65												80

Asn	Pro	Glu	Cys	Glu	Leu	Ile	Ser	Lys	Glu	Ser	Trp	Ser	Tyr	Ile	
			85												95

Val	Glu	Thr	Pro	Asn	Pro	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	Tyr	Phe
			100						105						110

Ala	Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe
			115						120						125

Glu	Arg	Phe	Glu	Ile	Phe	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr
			130						135						140

Val	Thr	Gly	Val	Ser	Ala	Ser	Cys	Ser	His	Asn	Gly	Lys	Ser	Ser	Phe
			145												160

Tyr	Arg	Asn	Leu	Leu	Trp	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn
			165						170						175

-continued

Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu
 180 185 190
 Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr
 195 200 205
 His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
 210 215 220
 Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu
 225 230 235 240
 Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
 245 250 255
 Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala
 260 265 270
 Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
 275 280 285
 Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300
 Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320
 Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335
 Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
 340 345 350
 Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365
 His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
 370 375 380
 Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
 385 390 395 400
 Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415
 Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
 420 425 430
 Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
 435 440 445
 Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
 450 455 460
 Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
 465 470 475 480
 Phe Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys
 485 490 495
 Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
 500 505 510
 Arg Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln
 515 520 525
 Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val
 530 535 540
 Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln
 545 550 555 560
 Cys Arg Ile Cys Ile
 565

<210> SEQ ID NO 9
 <211> LENGTH: 1695
 <212> TYPE: DNA

-continued

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 9

gatacagatc	ctgcactgca	ggctgccgtt	ggagcacatc	caaaaggaga	tggcgccag	60
agacaccagc	agcaccagag	aagaggccac	tgtgctgttag	atggccagga	tctgatacac	120
gccccatgctc	tccagttca	cggccatcgat	cttctcccg	ttcagcttgc	tctcctcgct	180
gtacttaggg	tagtctgttag	tgccgttctt	cacgctctcc	atacactcgt	tgttacactt	240
gtggtagaaac	tcgaaggcgc	cgttgccgat	ctccctggcg	ttgttcttca	gctggcttct	300
cacccattctca	tacaggttct	tcacggttgc	gtcgtggaa	tccagggtcc	tctcattctc	360
gaggaggacc	aggagttcg	cattgttagt	ccagatgtcc	agaaagccgt	cgtccacctt	420
cttgggttccgg	ttctccatcc	tccgctccag	cttggtaa	tccttgccca	cagcggtaaa	480
ctgggtgttc	atcttctcg	tcacgctgtt	cacccattgc	gtatggcg	tgtatggcg	540
ctgggttagac	ttctgtatcg	cggcatatcc	agagccctgc	tcatttctgtt	ggtggtagcc	600
gtaccagcca	tccaccatgc	ctgtccatcc	tcctctcgat	aatccggcga	tggctccaaa	660
cagtccctcg	ctctggatgc	tagggatgtt	tctcaggccc	gtcaccatcc	tcagcttggc	720
gtttctcaca	tacttggggc	actcgccgat	ggtcacagg	tgcacattct	ggaagggcag	780
gtgtcttattg	atggccccc	gagggtgtctg	gcacttggca	tcacactcat	ccatggggc	840
gttgcttgtg	atgatgccc	tgccaaagcc	tctgctcagg	gcaaaggcat	accaaggggc	900
gtcagattgc	ccgttggcct	cgaagatgtat	ggtatcgcca	ggctccagca	gggtccagta	960
gttaattgtatc	cggccctcc	ggctctctac	tttgggtctc	ttggcgatct	cgggggtgaa	1020
tcttctgtgt	tagtggctgg	acaccacgct	cacataggcg	ttctctgtgt	ggtacaggc	1080
ccgctgtattt	ccgatgttgg	gagggtgggt	cactccccc	agcaccagca	cttcctttc	1140
cttgggttcc	acgttagctct	tgctcagg	gggtacagg	ccatttctgc	ctgtcagcca	1200
cagcagggttc	cggttaga	tgcttttgc	gttgtggcta	cagctggcag	acacgcctgt	1260
cactgtgtga	ttagggcage	tggactcett	gggaaagatc	tcaaatctct	cgaagctgga	1320
cacgcttagac	agctgtcgc	gcagctcctc	gtaatcgcc	aagttagccag	ggtagcaggt	1380
gcatttc	ggattgggg	tctccacat	gtagctccag	cttcctttag	aaatcagcag	1440
ctcacactcg	gggttgc	gaatccatcc	ggccacagaa	caattgccc	gctgcagagg	1500
ggcaatgcct	ttcagcagac	acagcttgc	attgtggctg	tcctccagca	ggttcacaga	1560
gtgggtc	acgt	tctccagac	tgtatccac	gtgtcggtc	tattgttgc	1620
gtggtagccg	atacagattt	tgtcgccgt	ggtggcggt	aaggtacaca	gcagcaccag	1680
cagtttggcc	ttcat					1695

<210> SEQ ID NO 10

<211> LENGTH: 1551

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 10

atgaaggcca	aactgctgg	gctgctgtgt	acctttaccc	ccacccatgc	cgacacaatc	60
tgtatcggt	accacgcca	caatagcacc	gacaccgtgg	atacagtgt	ggagaagaac	120
gtgaccgtga	cccaactctgt	gaacctgtcg	gaggacagcc	acaatggcaa	gctgtgtctg	180
ctgaaaggca	ttgcccctct	gcagctgggc	aattgttctg	tggccggatg	gattctggc	240
aaccccgagt	gtgagctgt	gatttctaag	gagagctgga	gctacatcgt	ggagaccccc	300

-continued

aatcctgaga atggcacctg ctaccctggc tacttcgcgg attacgagga gctgcgcgag	360
cagctgtcta gcgtgtccag cttcgagaga ttcgagatct tccccaaaggga gtccagctgg	420
cctaattcaca cagtgacagg cgtgtctgcc agctgttagcc acaacggcaa aagcagcttc	480
taccggAACC tgctgtggct gacaggcaag aatggcctgt accccaaacct gagcaagagc	540
tacgtgaaca acaaggaaaa ggaagtgtcg gtgctgtggg gagtgcacca ccctccaaac	600
atcggaaatc agcgggcccgt gtaccacaca gagaacgcct atgtgagcgt ggtgtccagc	660
cactacagca gaagattcac ccccgagatc gccaagagac ccaaagttag agaccaggag	720
ggccggatca attactactg gaccctgtcg gagcctggcc ataccatcat cttecgaggcc	780
aacggcaatc tgatcgcccc ttggtatgcc tttggcctga gcagaggctt tggcagcggc	840
atcatcaca gcaacgcccc catggatgag tgtgatgcac agtgccagac acctcagggc	900
gccccatcaata gcagectgcc cttccagaat gtgcaccctg tgaccatcg cgagtgcggc	960
aagtatgtga gaagcggccaa gctgagaatg gtgaccggcc tgagaaacat ccctagcatc	1020
cagagcagag gactgtttgg agccatcgcc ggattcatcg agggaggatg gacaggcatg	1080
gtggatggct ggtacggcta ccaccaccag aatgagcagg gctctggata tgccggccat	1140
cagaagtcta occagaacgc catcaacggc atcaccaaca aggtgaacag cgtgatcgag	1200
aagatgaaca cccagttac cgctgtgggc aaggagttca acaagctgga gcggaggatg	1260
gagaacctga acaagaaggt ggacgacggc tttctggaca tctggaccta caatggcga	1320
ctcctggtcc tcctcgagaa tgagaggacc ctggacttcc acgacacgaa cgtgaagaac	1380
ctgtatgaga aggtgaagag ccagctgaag aacaacgcca aggagatcg caacggctgc	1440
ttcgagttt accacaagtg taacaacgag tttatggaga gctgtgaagaa cggcacctac	1500
gactacccta agtacagcga ggagagcaag ctgaaccggg agaagatcga t	1551

<210> SEQ ID NO 11

<211> LENGTH: 517

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 11

Met Lys Ala Lys Leu Leu Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr			
1	5	10	15

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr		
20	25	30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn		
35	40	45

Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile		
50	55	60

Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly			
65	70	75	80

Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile		
85	90	95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe		
100	105	110

Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe		
115	120	125

Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr		
130	135	140

Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe	
---	--

US 9,441,019 B2

79

80

-continued

145	150	155	160
Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn			
165	170	175	
Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu			
180	185	190	
Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr			
195	200	205	
His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg			
210	215	220	
Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu			
225	230	235	240
Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile			
245	250	255	
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala			
260	265	270	
Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met			
275	280	285	
Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser			
290	295	300	
Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro			
305	310	315	320
Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn			
325	330	335	
Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe			
340	345	350	
Ile Glu Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His			
355	360	365	
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr			
370	375	380	
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu			
385	390	395	400
Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu			
405	410	415	
Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu			
420	425	430	
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu			
435	440	445	
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys			
450	455	460	
Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys			
465	470	475	480
Phe Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys			
485	490	495	
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn			
500	505	510	
Arg Glu Lys Ile Asp			
515			

<210> SEQ ID NO 12
<211> LENGTH: 1551
<212> TYPE: DNA
<213> ORGANISM: Influenza virus
<400> SEQUENCE: 12

-continued

<210> SEQ ID NO 13
<211> LENGTH: 1554
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 13

atgaaggccta tcctggtggt gctgctgtac accttcgcca ccgccaacgc cgacaccctg 60
tgtcatcgctt accacgccaa caaacagcacc gacacctgtt ggagaagaac 120
gtgaccgtga cccacagcgtt gaacctgtctt gaggacaagg acaacggcaa gctgtcaag 180
ctggggggc tggccccctt gcacccggc aagtgcacaa tggccggctt gattctggc 240
aaccggagt gcgagagcctt gagcaccggc agcagctggt gctacatcgtt ggagacccc 300
agcagcgaca acggcacctt ctacccggc gacttcatcg actacgagga gctgcggag 360
cagctgagca gctgtggcag ctgcggcgg ttcgagatctt tcccaagac cagcagctgg 420
cccaaccacg acagcaacaa gggcgtgacc gccgcctgccc acggcggc cggcaagagc 480
ttctacaaga acctgatctt gctggtaag aaggcaaca gctacccaa gctgagcaag 540
agctacatca acgacaaggc caaggaggtt ctgggtgtt ggggcatttccca ccacccacgc 600

-continued

accaggcggcc accagcagag cctgtaccag aacgcccaga cctacgttt cgtaggcagc 660
agccggtaca gcaagaaggta caagcccgag atcgccatcc ggcccaaggt gcgggaccag 720
gaggggccgga tgaactacta ctggaccctg gtggagcccc gcgacaagat cacttcag 780
gccaccggca acctgggttgt gccccggta gccttcgcca tggagcgaa cgccggcagc 840
ggcatcatca tcagcgacac ccccggtcac gactgcaaca ccacctgcca gacccccaag 900
ggcgccatca acaccagcct gcccttcag aacatccacc ccatcaccat cggcaagtgc 960
cccaagtacg tgaagagcac caagctgcgg ctggccaccg gcctgcggaa catccccagc 1020
atccagagcc ggggcctgtt cggcgccatc gccggcttca tcgaggggcg ctggaccggc 1080
atggtgtgacg gctggtacgg ctaccaccac cagaacgagc agggcagcg ctacggcc 1140
gacctgaaga gcacccagaa cgccatcgac gagatcacca acaagggtgaa cagcgtatc 1200
gagaagatga acacccagtt caccggctgt ggcaaggagt tcaaccacct ggagaagcg 1260
atcgagaacc tgaacaagaa ggtggacgac ggcttcctgg acatctggac ctacaacgccc 1320
gagctgctgg tgctgctggaa gaacgagcg accctggact accacgacag caacgtgaag 1380
aacctgtacg agaagggtgcg gagccagctg aagaacaacg ccaaggagat cggcaacggc 1440
tgcttcgagt tctaccacaa gtgcgacaac acctgcatgg agagcgtgaa gaacggcacc 1500
tacqactacc ccaaqgtacag cqggdqgqcc aqgctqaacc qggqgqgqat cqac 1554

<210> SEO ID NO 14

<211> LENGTH: 518

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 14

Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Ala Thr Ala Asn
1 5 10 15

Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
20 25 30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
35 40 45

Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu Arg Gly Val
50 55 60

Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp Ile Leu Gly
65 70 75 80

Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp Ser Tyr Ile
85 90 95

Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro Gly Asp Phe
 100 105 110

Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro Asn His Asp

Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly Ala Lys Ser

Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro

Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu Val Val

Lam-Tung-Chan-Ida-Wing-Wing-Dawn-Song-Thien-Gang-Alia-Lam-Chun-Chun-Gang-Lam

195 200 205

-continued

Tyr Gln Asn Ala Asp Thr Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser
210 215 220

Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln
225 230 235 240

Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys
245 250 255

Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe
260 265 270

Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro
275 280 285

Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn
290 295 300

Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys
305 310 315 320

Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg
325 330 335

Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
340 345 350

Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
355 360 365

His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser
370 375 380

Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile
385 390 395 400

Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His
405 410 415

Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe
420 425 430

Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
435 440 445

Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu
450 455 460

Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly
465 470 475 480

Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val
485 490 495

Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu
500 505 510

Asn Arg Glu Glu Ile Asp
515

<210> SEQ_ID NO 15
<211> LENGTH: 1554
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 15

gtcgatotcc tccccgttca gcttggcctc ctcgctgtac ttggggtagt cgttagtgcc 60
gttcttcacg ctctccatgc aggtgttgc gcacttgtgg tagaactcga agcagccgtt 120
gccgatctcc ttggcggtgt tcttcagctg gctccgcacc ttctcgtaa ggttcttcac 180
gttgctgtcg tggtagttcca gggtccgcctc gttctccagc agcaccagca gctcggcggtt 240
gttaggtccag atgtccagga agccgtcgctc caccttcttg ttcaggttct cgatccgctt 300

-continued

ctccagggtgg ttgaactcct tgcccacggc ggtgaactgg gtgttcatct tctcgatcac	360
gctgttacc ttgttgttga tctcgatcgat ggctgttctgg tgctttca ggtcgccggc	420
gtagccgctg ccctgtcgat tctgggtggt gtagccgtac cagccgtcca ccatgcccgt	480
ccagccgccc tcgatgaagc cggcgatggc gccgaacagg ccccgctct ggtatgtggg	540
gtatgttccgc aggccgggtgg ccagccgcag ctgggtgttc ttacgtact tggggcactt	600
ggcgatgttg atgggggttga tggttctggaa gggcaggctg tggttcatgg cgcccttggg	660
gggtctggcag gtgggtttgc agtctgtcac ggggggtgtcg ctgtatgtga tggcgatgcc	720
ggcgatgttccgc tccatggcga aggccgtaccg gggcaccacc aggttgcggg tggccctcgaa	780
ggtgatcttg tggccgggct ccaccagggt ccagttagtgc ttcatccggc cctctgtgtc	840
ccgcacatttggccggatgg cgatctcggtt ctggtaacttc ttgtgttacc ggctgtgtcc	900
cacgaacacg taggtgtcggtt cgttctggta caggctctgc tggtcgccgc tggtgttggg	960
gtgggtggatg cccccacagca ccagcacccctt ctggcccttg tggttcatgtt agctttgtct	1020
cagcttgggg tagctgttgc ccttcttcac cagccagatc aggttcttgc agaagctttt	1080
ggcgccggcg tggggggcagg cggcggtcac gcccgggttg ctgtcggtt tggtccagct	1140
gctgggttttgc gggaaagatct cgaaccgtctc gaagctgtctc acgctgtctca gctgtcccg	1200
cagcttctcg tagtctgtatgc agtctgttgc agtctgttgc cttgttgc tgctttttttt	1260
ctccacatgttgc tagctttccagc tgctggccgtt gctcaggttc tggcacttcgg gggttgc	1320
aatccagocgcg gctgtgttgc acttgcggcag gtgcagggggg gcccacccccc gcagcttgc	1380
cagcttgcggc ttgtgtttgtt cctccagccag gttcacgttgc tggttgcacgg tcaacgttctt	1440
ctccacatgttgc tagctttccagc tgctggccgtt gctcaggttc tggcacttcgg gggttgc	1500
gtgttccacgg tgctgggtgtt gttgttggcg tggttagccgc tgcacagggt	1554
gtcgccgggtt ggggtggcga aggtgtacag cagcaccacc accgttgcct tcat	1554

<210> SEQ ID NO 16
<211> LENGTH: 1542
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 16

atggccatca tctacccgtat cctgtgttt acagctgtgc gggcgatca gatctgtatc
ggcttaccacg ccaacaatag caccgagaag gtggacacca tcctggaaag aaatgtgacc 120
gtgacccacg ccaaggatata tctggaaaag acccacaacg gcaagctgtg caagctgaat
ggcattectc ctctggaaact gggcgattgt tctattgtg gctggctgct gggaaatcct 180
gagtgcgata gactgtgtc tggccttag tggagctaca tcatggaaaa agagaaccct 300
agggacggac tgggttaccc cgccgacccaa aacgattacg aggaactgaa gcacccgtctg 360
tccagcgtga agcacttcga gaaagtgaag atccctggcca aggatagatg gaccacgcat 420
acaacaacag gcccggaaacg agcttgcgtc gtgtccggca accccagctt cttcgaaaat 480
atggtctggc tgaccaagaa gggctctaata tattcctgtgg ccaagggcag ctacaataat 540
acaagcggcg agcagatgtt gattatttgg ggctgtgcacc accctaatga tgagacagag 600
cagagaaccc tggatccagaa tggggcaca tacgtgtctg tgggcaccag cacactgaat 660
aagagaagca cccccgatata tgccaccaga cccaaagtga atggacaggg cgccagaatg 720
gaattttctt ggacctgtctt ggatgtgg gacaccatca actttgagag caccggaaat 780
ctgattgccc ctgagttacgg cttcaagatc agcaagagag gcagcagcgg catcatgaaa 840
acagaggggca ccctggaaaa ctgtgaaacc aagtgtcaga cacctctggg cgccattaat 900

-continued

```

accacccctgc cttccataa tggcacccct ctgacaatcg gcgagtgccc taagtacgtg      960
aagtctgaga aactggtgct ggccacagga ctgagaaatg tgccccagat cgagtcaaga     1020
ggcctgtttg gagccattgc cggctttatt gaaggcggat ggcagggaat ggtggatgg      1080
tggtacggct atcaccacag caatgatcg gatatctggc atgccgcca taaagagagc     1140
acccagaagg ctttgacgg catcaccaac aaagtgaaca gcgtgatcga gaagatgaac     1200
acccagttt aggccgtggg caaagagttc agcaatctgg aaagacggct ggaaaacctg     1260
acaagaaaa tggaaagatgg cttctggac gtgtggacat ataatgccg gctgtggc      1320
ctgatggaaa acgagaggac cctggactt cacgacgaca acgtgaagaa cctgtacgac     1380
aaagtgcgga tgcagctgag agacaatgtg aaagagctgg gcaacggctg ctttgagttc     1440
taccacaagt gcgacgacga gtgcatgaat agcgtgaaga acggcaccta cgactaccct     1500
aagtatgagg aagagagcaa gctgaacaga aacgagatca ag                               1542

```

<210> SEQ ID NO 17

<211> LENGTH: 514

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 17

```

Met Ala Ile Ile Tyr Leu Ile Leu Leu Phe Thr Ala Val Arg Gly Asp
1           5           10          15

```

```

Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Lys Val Asp
20          25          30

```

```

Thr Ile Leu Glu Arg Asn Val Thr Val Thr His Ala Lys Asp Ile Leu
35          40          45

```

```

Glu Lys Thr His Asn Gly Lys Leu Cys Lys Leu Asn Gly Ile Pro Pro
50          55          60

```

```

Leu Glu Leu Gly Asp Cys Ser Ile Ala Gly Trp Leu Leu Gly Asn Pro
65          70          75          80

```

```

Glu Cys Asp Arg Leu Leu Ser Val Pro Glu Trp Ser Tyr Ile Met Glu
85          90          95

```

```

Lys Glu Asn Pro Arg Asp Gly Leu Cys Tyr Pro Gly Ser Phe Asn Asp
100         105         110

```

```

Tyr Glu Glu Leu Lys His Leu Leu Ser Ser Val Lys His Phe Glu Lys
115         120         125

```

```

Val Lys Ile Leu Pro Lys Asp Arg Trp Thr Gln His Thr Thr Gly
130         135         140

```

```

Gly Ser Arg Ala Cys Ala Val Ser Gly Asn Pro Ser Phe Phe Arg Asn
145         150         155         160

```

```

Met Val Trp Leu Thr Lys Lys Gly Ser Asn Tyr Pro Val Ala Lys Gly
165         170         175

```

```

Ser Tyr Asn Asn Thr Ser Gly Glu Gln Met Leu Ile Ile Trp Gly Val
180         185         190

```

```

His His Pro Asn Asp Glu Thr Glu Gln Arg Thr Leu Tyr Gln Asn Val
195         200         205

```

```

Gly Thr Tyr Val Ser Val Gly Thr Ser Thr Leu Asn Lys Arg Ser Thr
210         215         220

```

```

Pro Asp Ile Ala Thr Arg Pro Lys Val Asn Gly Gln Gly Arg Met
225         230         235         240

```

```

Glu Phe Ser Trp Thr Leu Leu Asp Met Trp Asp Thr Ile Asn Phe Glu
245         250         255

```

-continued

Ser Thr Gly Asn Leu Ile Ala Pro Glu Tyr Gly Phe Lys Ile Ser Lys
260 265 270

Arg Gly Ser Ser Gly Ile Met Lys Thr Glu Gly Thr Leu Glu Asn Cys
275 280 285

Glu Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Thr Thr Leu Pro
290 295 300

Phe His Asn Val His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr Val
305 310 315 320

Lys Ser Glu Lys Leu Val Leu Ala Thr Gly Leu Arg Asn Val Pro Gln
325 330 335

Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
340 345 350

Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Ser Asn
355 360 365

Asp Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala
370 375 380

Phe Asp Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn
385 390 395 400

Thr Gln Phe Glu Ala Val Gly Lys Glu Phe Ser Asn Leu Glu Arg Arg
405 410 415

Leu Glu Asn Leu Asn Lys Met Glu Asp Gly Phe Leu Asp Val Trp
420 425 430

Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr Leu
435 440 445

Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Met
450 455 460

Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe
465 470 475 480

Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly Thr
485 490 495

Tyr Asp Tyr Pro Lys Tyr Glu Glu Ser Lys Leu Asn Arg Asn Glu
500 505 510

Ile Lys

<210> SEQ_ID NO 18
<211> LENGTH: 1542
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 18

cttgatctcg	tttctgttca	gcttgctc	ttectcatac	ttagggtagt	cgtaggtgcc	60
gttcttcacg	ctattcatgc	actcgctgc	gcacttgtgg	tagaactcaa	agcagccgtt	120
geccagctct	ttcacattgt	ctctcagctg	catccgcact	ttgtcgta	cggttcttac	180
gttgctgtcg	tgaaagtcca	gggtcctc	gtttccatc	agcaccagca	gctcgccatt	240
atatgtccac	acgtccagga	agccatctc	cattttctt	ttaggtttt	ccagccgtct	300
ttccagattg	ctgaactctt	tgcccacggc	ctcaaactgg	gtgttcatct	tctcgatcac	360
gctgttca	ctgttgttga	tgccgtcaaa	ggccttctgg	gtgctcttt	tatcgccggc	420
atagccagat	ccctgatcat	tgctgtgg	atagccgtac	cacccatcca	ccattccctg	480
ccatccgcct	tcaataaagc	cgccaatggc	tccaaacagg	ccttttgact	cgatctgggg	540
cacatttctc	agtccgtgg	ccagcaccag	tttctcagac	ttcacgtact	tagggactc	600
gcccattgtc	agagggtgca	cattatggaa	gggcagggtg	gtattaatgg	cggccagagg	660

-continued

tgtctgacac ttggttcac agtttccag ggtgcctct gtttcatga tgccgctgct	720
gctctcttg ctgatcttga agccgtactc aggggcaatc agattccgg tgctctcaaa	780
gttgatggtg tccccatata ccagcagggt ccaggaaaat tccattctgc cgccctgtcc	840
atccactttg ggtctggtgg caaatatcggtt ggtgcctctc ttattcagtg tgctggtgcc	900
cacagacacg tatgtgccc cattctggta cagggttctc tgctctgtct catcattagg	960
gtgggtgcacg ccccaataa tcagcatcg ctgcgcgtt gtattattgt agctgccctt	1020
ggccacagga taatttagagc cttcttttgtt cagccagacc atatttctga agaagctggg	1080
gttgcggac acagcacaag ctctgccttc gcctgttgtt gtatgtggg tccatctatc	1140
cttggcagg atcttcaattt tctcgaagtgc ttccacgtt gacagcagggt gcttcagttc	1200
ctcgtaatcg ttgaagctgc cggggtaaca cagtccgtcc cttagggttct cttttccat	1260
gatgtagctc cactcaggca cagacagcag tctatcgac tcaggatttc ccagcagcca	1320
gcacagcaata gaacaatcgc ccagttccag aggaggaatg ccattcagct tgcacagctt	1380
gcgcgttgtgg gtctttcca gaatatcctt ggctgtggtc acggtcacat ttctttccag	1440
gatgggtgtcc accttctcggt tgcttattgtt ggctgtggtag ccgatacaga tctgtatcgcc	1500
ccgcacagct gtaaaacagca ggatcaggta gatgatggcc at	1542

<210> SEQ ID NO 19

<211> LENGTH: 1557

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 19

atggaaaacca tcattgcct gagctacatc ttttgtctgg ctctggcca ggatctgcc	60
ggcaatgata atageacccgc caccctgtgt ctgggacacc acgcccgtgcc taatggcacc	120
ctgggtaaaa ccattaccga cgaccagatc gaagtgcaccatg ccaccggcgtcc gctgggtcag	180
agcagcagca cccgcacatcg ctgcacaaac ccccacagaa tcctggatgg catcgactgt	240
accctgtatcg atgcctgtctt gggggatctt cactgcgcac tggtccagaa cgagacatgg	300
gacctgttctg tggagagaag caaggcccttc agcaactgctt acccctacga tgtgcggat	360
tacgcctctc tgagaaggctt ggtggccagc agcggcacac tggaattcat caccgaggc	420
tttaccttggc cagggctgtac ccagaatggc ggcagcaatg cctgtaaaag aggccttggc	480
agcggcttctt tcagcagactt gaaactggctt accaagtccgc gcagcaccta ccctgtgt	540
aacgtacca tgcccaacaa cgacaacttc gacaagctgtt acatctgggg cgtgcaccac	600
ccttagcacca atcaggaaca gaccagcctg tacgtgcagg ccagcggcag agtgcacgt	660
tctaccagac ggtcccagca gaccatcatc cccacatcg agtcaagaccc ttgggtgcgc	720
ggcctgagca gcagaatcatcg catctactgg accatcgatgc aacctggcga cgtgttgt	780
atcaacagca atggcaaccc gatcgcccc agaggctact tcaagatgcg gaccggcaag	840
agcagcatca tgagaaggcga cggcccccattt gatacctgttac tcaagcgatgtt catcaccccc	900
aacggcgcacca tccccaaacga caaggcccttc cagaacgttca acaagatcac ctacggcgcc	960
tgccttaagt acgtgaagca gaacaccctgtt aagctggcca cggcatgag aaatgtgcc	1020
gagaaggcaga caagaggcctt gtttggccatcc attggccgtt ttatcgagaa cggctggag	1080
ggcatgtatcg atgggtggta cggcttcaga caccagaattt ctgaggggcac aggacaggcc	1140
ccgcatgtatcg agtctacaca ggccgcattt gaccatgttca acggcaagctt gaacagatgtt	1200

US 9,441,019 B2

95**96**

-continued

```

atcgagaaaa ccaacgagaa gttccaccag atcgagaaaag aattcagcga ggtggaggc 1260
agaatccagg acctggaaaa atacgtggag gacaccaaga tcgacctgtg gagctacaat 1320
gccgaactgc tggtegcctt ggaaaaccag cacaccatcg acctgaccga cagcgagatg 1380
aataagctgt tcgaaaagac cagacggcag ctgagagaaa acgcccggaga catgggcaac 1440
ggctgcttca agatctacca caagtgcgac aacgcctgca tcgagagcat cagaaacggc 1500
acctaaccacc accatgtgtta cagggacgag gccctgaaca acagattcca gatcaag 1557

```

<210> SEQ ID NO 20

<211> LENGTH: 519

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 20

```

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Phe Cys Leu Ala Leu Gly
1           5          10          15

```

```

Gln Asp Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20          25          30

```

```

His His Ala Val Pro Asn Gly Thr Leu Val Lys Thr Ile Thr Asp Asp
35          40          45

```

```

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
50          55          60

```

```

Gly Lys Ile Cys Asn Asn Pro His Arg Ile Leu Asp Gly Ile Asp Cys
65          70          75          80

```

```

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro His Cys Asp Val Phe Gln
85          90          95

```

```

Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Phe Ser Asn
100         105         110

```

```

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
115         120         125

```

```

Ala Ser Ser Gly Thr Leu Glu Phe Ile Thr Glu Gly Phe Thr Trp Thr
130         135         140

```

```

Gly Val Thr Gln Asn Gly Gly Ser Asn Ala Cys Lys Arg Gly Pro Gly
145         150         155         160

```

```

Ser Gly Phe Phe Ser Arg Leu Asn Trp Leu Thr Lys Ser Gly Ser Thr
165         170         175

```

```

Tyr Pro Val Leu Asn Val Thr Met Pro Asn Asn Asp Asn Phe Asp Lys
180         185         190

```

```

Leu Tyr Ile Trp Gly Val His His Pro Ser Thr Asn Gln Glu Gln Thr
195         200         205

```

```

Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr Val Ser Thr Arg Arg
210         215         220

```

```

Ser Gln Gln Thr Ile Ile Pro Asn Ile Glu Ser Arg Pro Trp Val Arg
225         230         235         240

```

```

Gly Leu Ser Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
245         250         255

```

```

Asp Val Leu Val Ile Asn Ser Asn Gly Asn Leu Ile Ala Pro Arg Gly
260         265         270

```

```

Tyr Phe Lys Met Arg Thr Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
275         280         285

```

```

Pro Ile Asp Thr Cys Ile Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
290         295         300

```

```

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Lys Ile Thr Tyr Gly Ala
305         310         315         320

```

-continued

Cys Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
325 330 335

Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Leu Phe Gly Ala Ile Ala
340 345 350

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Ile Asp Gly Trp Tyr Gly
355 360 365

Phe Arg His Gln Asn Ser Glu Gly Thr Gly Gln Ala Ala Asp Leu Lys
370 375 380

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Val
385 390 395 400

Ile Glu Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
405 410 415

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
420 425 430

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
435 440 445

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
450 455 460

Glu Lys Thr Arg Arg Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
465 470 475 480

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Glu Ser
485 490 495

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
500 505 510

Asn Asn Arg Phe Gln Ile Lys
515

<210> SEQ ID NO 21
<211> LENGTH: 1557
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 21

```

cttgatctgg aatctgttgc tcagggcctc gtccctgtac acatcggtt cgttaggtgcc      60
gtttctgtat ctctcgatgc aggcggtgtc gcacttgtgg tagatcttga agcagccgtt     120
gecccatgtcc tcggcggttt ctctcagtcg ccgtctggtc ttttcaaca gcttattcat     180
ctcgctgtcg gtcaggtcga tgggtgtgtc gtttccagg gcgaccagca gttcggcatt     240
gttagctccac aggtcgatct tgggtgtcctc cacgtatttt tccagggtct ggattctgcc    300
ctccacactcg ctgaattctt tctcgatctg gtggaaacttc tctgtgggtt tctcgatcac    360
tctgttcagc ttggcggtga tctgggtcgat ggccggcctgt gtagacttca gatcgccggc    420
ctgtcctgtg ccctcagaat tctgggtgtct gaagccgtac cacccatcga tcatgcctc    480
ccagccgttc tcgataaaggc cggcaatggc gccaaacagg cctcttgcgtt gcttcgggg    540
cacatttctc atgcccgtgg ccagcttcag ggtgttctgc ttcacgtact tagggcaggc    600
gccgttaggt atcttggta cgttctggaa gggcttgcgtc ttggggatgc tgccgttggg    660
ggtgatgcac tcgctgatac aggtatcgat gggggcgctcg cttctcatga tgctgcttt    720
gccggccgc atcttgaagt agcctctggg ggcgatcagg ttgccattgc tggtgatcac    780
cagcacgtcg ccagggttca cgtgggtcca gtagatgcgtt attctgtgtc tcaggccgcg    840
cacccaaggt cttgactcga tggggggat gatggtctgc tgggaccgtc tggtagacac    900
ggtcactctg ccgctggcct gcacgtacag gctggctgtc tctgtattgg tgcttagggtg   960

```

US 9,441,019 B2

99**100**

-continued

```

gtgcacgccc cagatgtaca gcttgcgaa gttgtcggt tgggcatgg tcacgttcag    1020
cacagggttag gtgctgcgg acttggtcag ccagttcagt ctgtgaaga agccgctgcc    1080
agggcctt ttacaggcat tgctgccgg attctgggtc acgcctgtcc aggtaaagcc    1140
ctcgggtatg aattccagtg tgccgctgtc ggccaccagg ctttcagag aggcttaatc    1200
gggcacatcg taggggttagc agttgcgtaa ggccgttgcgtt ctctccacga acaggtccca    1260
tgtctcggtc tggAACACGT cgccgtgagg atcgcccagg agggcatcga tcagggtaca    1320
gtcgatgcca tccaggattc tgtgggggtt gttgcagata ttgcgggtgc tgctgtctg    1380
caccagctcg gtggcattgg tcacttcgt ctggcgtcgt gtaatgggtt tcaccagggt    1440
gecattaggc acggcgtggt gtcccagaca cagggtggcg gtgcatttat cattgccggg    1500
cagatccctgg occagagcca gacaaaagat gtgcgtcagg gcaatgtgg ttttcat    1557

```

<210> SEQ ID NO 22
<211> LENGTH: 1557
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 22

```

atgaaaaacca tcattgcct gagctacatc ctgtgcgtgg tgttcacaca gaagctgccc    60
ggcaacgata atagcaccgc cacactgtgt ctggacacc acgcgtgcc taatggcacc    120
atcgtaaaa caatcaccaa cgaccagatc gaagtgacca atgccacaga gctggcgtcag    180
agcagcagca cagggcagat ctgtgacagc ccccaccaga ttctggatgg cgagaactgt    240
accctgtatcg atgcctgtc gggegatct cagtgcgtc gttccagaa caagaatgg    300
gacctgttcg tggagagaag caaggcctac agcaactgtc acccctacga cgtgcctgtat    360
tacgcccagcc tgagaaggctt ggtggcctt agcggcaccc tggaaattcaa caacgagac    420
ttcaactgga cggcgtgac acagaatggc accagcagcg cctgcgtcag acggcctaac    480
aacagttct tcaatgtact gatgggtcg acccactgtc agttcaagta cccgcctg    540
aacgtgacca tgcccaacaa tgagaagttc gacaagctgt acatctgggg agtgcaccac    600
cctggcaccg acaacgatca gatcttcct tacgcccagg ccagcggcag aatcaccgtg    660
tccaccaaga gaagecagca gacgtgtc cccatatcg gcagcagacc cagagtgcgg    720
aacatccccca gcaggatcg catctactgg acaatgtgtc agcctggcgtc catctgtg    780
atcaacgca cggcgtaccc gatcgccctt cggggctact ttaatgtcag aagcggcaag    840
agcagcatca tgagatccga cggccatc ggcaagtgc acagcgtc gatcaccacca    900
aacggcagca tcccaacga caaggccctt cagaacgtgtc acaggatcac ctacggcgcc    960
tgcctatgt acgtgaaagca gaacaccctg aagctggcca cggcatgtc aaatgtgccc    1020
gagaaggcaga ccagaggcat cttggcgtcc attgcccgtt ttagtgcgtt tggctggag    1080
ggaatgggtgg atgggtggta cggcttcaga caccagaata gcgaggaaat tggacaggcc    1140
ggcgatctga aatctaccca ggccgtccatc gaccagatca acggcaagct gaacaggctg    1200
atcggtcaaga ocaacgagaa gttccaccag atcgagaaat aattcagcgtc ggtggaggc    1260
agaatccagg acctggaaaa atacgtggag gacaccaaga tcgacgtgtc gagctacat    1320
ggcgaaactgc tggcgtccatc ggaaaaccag cacacaattt atctgcgtc cagtgcgtc    1380
aataagctgt tcgagaaaaac caagaagcgtc ctgagagaaa acggcgtggaaat catggcgtc    1440
ggctgtttca agatctaccca caagtgcgtc aacgcgtgtc tcggcgtcgtc cagaaacggc    1500

```

-continued

acctacgacc acgacgtgta cagagatgag gccctgaaca accggttca gatcaag 1557

<210> SEQ ID NO 23
<211> LENGTH: 519
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 23

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Thr
1 5 10 15

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20 25 30

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Asn Asp
35 40 45

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
50 55 60

Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Glu Asn Cys
65 70 75 80

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
85 90 95

Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
100 105 110

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
115 120 125

Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
130 135 140

Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Asn
145 150 155 160

Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Lys Phe Lys
165 170 175

Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys
180 185 190

Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Asn Asp Gln Ile
195 200 205

Phe Pro Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
210 215 220

Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
225 230 235 240

Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
245 250 255

Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
260 265 270

Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
275 280 285

Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
290 295 300

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
305 310 315 320

Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
325 330 335

Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala
340 345 350

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
355 360 365

-continued

Phe Arg His Gln Asn Ser Glu Gly Ile Gly Gln Ala Ala Asp Leu Lys
 370 375 380

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu
 385 390 395 400

Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
 405 410 415

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
 420 425 430

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
 435 440 445

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
 450 455 460

Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
 465 470 475 480

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
 485 490 495

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
 500 505 510

Asn Asn Arg Phe Gln Ile Lys
 515

<210> SEQ_ID NO 24

<211> LENGTH: 1557

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 24

```

cttgatctga aaccgggttgt tcagggcctc atctctgtac acgtcggttgt cgtaggtgcc       60
gtttctgtatg ctgcccgtatgc aggcgttgc gcacttggtagatcttgc agcagccgtt       120
gccccatgtcc tcggcggtttt ctctcagctg cttcttgggtt ttctcgaaca gcttattcat       180
ctcaactgtct gtcaagatcaa ttgtgtgtctg gtttccagg gcgaccagca gttcggcatt       240
gtagctccac aggtcgatct tgggtgcctc cacgtatttt tccaggtctt ggattctgcc       300
ctcccacctcg ctgaattctt tctcgatctg gtggaaacttc tctgtggctt tgccgatctag       360
ctgttccagc ttggcggttga tctgggtcgat ggccggctgg gtagatttca gatccggcgc       420
ctgtccaatt ccctcgctat tctgtgtctt gaagccgtac cacccatcca ccattccctc       480
ccagccattc tcgataaaagc cggcaatggc gccaaagatg cctctggctt gcttctcggg       540
cacatttctc atgcccgggtgg ccagttcag ggtgttctgc ttcacgtatc tagggcaggc       600
gecgtaggtg atcctgttca cgttctggaa gggcttgc gtagatgtcg tgccgttgg       660
ggtagatgcac tctgtgttgc acttgcgtat gggggcgctcg gatctcatga tgctgtcttt       720
gecgcttctg atcttaaagt agccccggagg ggcgatcagg ttgcccggtgc tggtgatcag       780
caggatgtcg ccagggttca cgattgttca gtagatgtcg atcctgtctt ggtgttccg       840
cactctgggt ctgctgccga tattggggat cacggtctgc tggcttctt tggtggacac       900
ggtagtgcac tccgtggcct gggcgtaagg gaagatctga tctgtgtcg tgccagggtg       960
gtgcactccc cagatgtaca gcttgcgaa cttctcattt tgccatgg tcacgttctc       1020
ggcgccccgtac ttgaacttca ggtgggtcag ccaattcagt ctactgaaga agctgttgc       1080
ggaccgtctg atgcaggcgc tctgtgtgc attctgttgc acggccgtcc agtggaaatc       1140
ctcggttgc aattccagg tgccgttca ggcaccagg cttctcaggc tggcgtaatc       1200
aggcacgtcg tagggtagc agttgttgc ggccttgcctt ctctccacga acagggtccca       1260

```

-continued

```

tttcttgttc tggaagccgt cgcaactgagg atcgcccagc agggcatcga tcagggtaca 1320
gttctcgcca tccaggatct ggtggggct gtcacagatc tcgcctgtc tgctgtctg 1380
caccagctct gtggcattgg tcacttcgtat ctggtcgttg gtgattgtt tcacgtatgg 1440
gccatttaggc acggcgtggt gtcccagaca cagtgtggcg gtgcattttat cgttgccggg 1500
cagttctgt gtgaacacca ggcacaggat gtagctcagg gcaatgtatgg ttttcat 1557

```

<210> SEQ ID NO 25

<211> LENGTH: 1560

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 25

```

atggaaaaga tcgtgtgtct gctggccatt gtgagccctgg tgaagagcga ccagatctgc 60
attggctacc acgccaacaa tagcacagag caggtggaca ccatcatggaa aaaaaacgtg 120
accgtgaccc acgctcagga catctggaa aagaccacaa acggcaagct gtgtgtatctg 180
gacggcgtga agcctctgtat cctgagagat tgttagcgtgg ctggatggct gctgggcaac 240
cctatgtgcg acgagttcat caacgtgccc gagtggagct atatgtggaa gaaggccaaac 300
cccaccaacg atctgtgtta cccoggcagc ttcaacgatt acgaggaact gaagcacctg 360
ctgtccccgaa tcaaccactt cgagaagatc cagatcatcc ccaagtcctc ttggagcgtat 420
cacgaagcct ctageggagt gtctagcgtcc tgccttacc tggcagccc cagtttttc 480
agaaacgtgg tgtggctgtat caagaagaac agcacctacc ccaccatcaa gaagagctac 540
aacaacacca accaggaaga tctgctggtc ctgtggggaa tccaccaccc taatgtatgcc 600
gccgagcaga ccagactgtaa ccagaaccccc accacctata tcagcatcgg caccagcacc 660
ctgaatcaga gactggtgcc caagatcgtcc accagatcca aggtgaacgg ccagagcggc 720
aggatggaat tcttgtggac catctgtaa cccaaacgcg ccatcaactt cgagagcaac 780
ggcaacttta tcgccccctga gtaagccctac aagatcgtga agaagggcga cagcgcctac 840
atgaagagcg agctggaaata cggcaactgc aacaccaagt gccagacacc tatggggcgc 900
atcaacagca gcatgcccatt ccacaacatc caccctctga ccatcggcga gtgccttaag 960
tacgtgaaga gcaacagact ggtgtggcc acaggcctga gaaatagccc ccagegggg 1020
agcagaagaa agaagagggg cctgtttggaa gccatcgtcc gctttattga aggccggctgg 1080
caggaaatgg tggatggctg gtacggctac caccacagca atgagcaggg ctctggatata 1140
gccgcccaca aagagtctac ccagaaggcc atcgacggcg tcaccaacaa ggtgaacagc 1200
atcatcgtaca agatgaacac ccagttcgag gctgtggcc gagagttcaa caacctggaa 1260
cgccggatcg agaacctgaa caagaaaatg gaagatggct tcctggatgt gtggacctac 1320
aatgcccgaac tgctgggtct gatggaaaac gagcggacc tggacttcca cgacagcaac 1380
gtgaagaacc tgcgtacgacaa agtgcggctg cagctgagag acaacgccaa agagctggc 1440
aacggctgct tcgagttcta ccacaagtgc gacaacgagt gcatggaaag catcaggaaac 1500
ggcacctaca actacccctca gtacagcgtggc gaagccaggc tgaagaggaa agagatcagc 1560

```

<210> SEQ ID NO 26

<211> LENGTH: 520

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 26

-continued

Met Glu Lys Ile Val Leu Leu Ala Ile Val Ser Leu Val Lys Ser
 1 5 10 15
 Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 20 25 30
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 35 40 45
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
 50 55 60
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
 65 70 75 80
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
 85 90 95
 Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
 100 105 110
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
 115 120 125
 Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
 130 135 140
 Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
 145 150 155 160
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
 165 170 175
 Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
 180 185 190
 Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
 195 200 205
 Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
 210 215 220
 Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
 225 230 235 240
 Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
 245 250 255
 Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
 260 265 270
 Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
 275 280 285
 Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
 290 295 300
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
 305 310 315 320
 Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
 325 330 335
 Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
 340 345 350
 Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
 355 360 365
 Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
 370 375 380
 Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
 385 390 395 400
 Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
 405 410 415
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp

-continued

420	425	430	
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu	Leu Leu Val	Leu Met	
435	440	445	
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser	Asn Val Lys	Asn Leu	
450	455	460	
Tyr Asp Lys Val Arg Leu Gln	Leu Arg Asp	Asn Ala Lys Glu Leu Gly	
465	470	475	480
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp	Asn Glu Cys Met	Glu	
485	490	495	
Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln	Tyr Ser Glu	Glu Ala	
500	505	510	
Arg Leu Lys Arg Glu Glu Ile Ser			
515	520		

<210> SEQ_ID NO 27
<211> LENGTH: 1560
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 27

gtgtatcttccctttca	gcctggcttc	ctcgctgtac	tgagggttagt	tgttaggtgcc	60
gttcctgatgtttccatgc	actcggttgc	gcacttgtgg	tagaactcga	agcagccgtt	120
gcccagctcttggcggtgt	ctctcagctg	cagccgact	ttgtcgta	ggttttcac	180
gttgctgtcg	tggaagtcca	gggtccgctc	gtttccatc	agcaccagca	240
gttaggtccac	acatccagga	agccatcttc	cattttcttg	ttcaggttct	300
ttccaggttg	tgtaactctc	tgcccacagc	ctcgaactgg	gtgttcatct	360
gtgttccacc	ttgttgtga	cgccgtcgt	ggccttctgg	gttagacttt	420
atatccagag	ccctgctcat	tgctgtggtg	gtagccgtac	cagccatcca	480
ccagccgcct	tcaataaaagc	cggcgatggc	tccaaacagg	ccccctttct	540
ctccccgtgg	gggtctatttc	tcaggcctgt	ggccagcacc	agtctgttc	600
cttagggcac	tcgcccgtgg	tcagagggtg	gtatgttggg	aagggcatgc	660
ggcgccdata	ggtgtctggc	acttgggttt	geagttggcc	tattccagct	720
gatggcgctg	tcgccccttc	tcacatctt	gtaggcgta	tcagggggca	780
gttgctctcg	aagttgtatgg	cgtcggtgg	cttcaggatg	gtccagaaga	840
gcccgtctgg	ccgttcaccc	tggatctgtt	ggcgtatcttgc	tctgttccat	900
ggtgtctgg	ccgatgctga	tataggtgg	ggggttctgg	tacagtctgg	960
ggcatcatta	gggtgggtgg	ttccccacag	gaccagcaga	tcttcctgg	1020
gtagcttcc	ttgatgggtgg	ggtaggtgtct	gttcttcttg	atcagccaca	1080
gaagaagctg	gggctgcccc	ggtaaggaca	ggcgctagac	actccgttag	1140
atcgctccaa	gaggacttgg	ggatgtatcg	gatcttctcg	aagtgggtga	1200
caggtgttcc	atttccctcg	aatcggttga	gctgccccgg	taacacagat	1260
gttggcccttc	tccacatata	agctccactc	gggcacgttg	atgaactcgt	1320
gttgcccagc	agccatccag	ccacgctaca	atctctcagg	atcagaggct	1380
cagatcacac	agcttgcgt	tgtgggtctt	ttccaggatg	tcttgacgct	1440
cacgtttttt	tccatgtatgg	tgtccacctg	ctctgtgtca	ttgttggcgt	1500
gcagatctgg	tcgctttca	ccaggctcac	aatggccagc	agcagcacga	1560

-continued

<210> SEQ ID NO 28
<211> LENGTH: 1602
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

>400 > SEQUENCE: 28

atgaaggcca tcategtgct gctgatggtg gtgaccagca acggccatag aatctgcacc 60
ggcatcacca gcagcaatacg cccccatgtg gtaaaaacag ccaccaggcg cgaagtgaat 120
gtgacaggcg tgatccctct gaccaccacc cccaccaaga gctacttcgc caacctgaag 180
ggcaccagaa ccagaggcaa gctgtcccc gattgcctga actgcaccga tctggatgtg 240
gctctggca gacatatgtg tggggcacc acaccatgtg ccaaggccag catctgcac 300
gaagtgaagc ctgtgaccag cggctgcctt cccatcatgc acgaccggac caagatcaga 360
cagctgccca acctgtctgag aggctacgag aacatccggc tgtccaccca gaatgtgatc 420
gatgccgaga aagccccctgg cggaccttat agactggca ccagggctc ttgtcccaat 480
gccacccctca agagcggctt ttttgcaca atggcctggg ccgtgcctaa ggacaacaac 540
aagaacgcca ccaaccctct gaccgtggag gtgccttaca tctgtacaga gggcgaggat 600
cagatcacag tgtggggctt ccacagcgcac gacaagaccc agatgaagaa cctgtacggc 660
gacagcaacc cccagaagtt taccagcagc gccaatggcg tgaccaccca ctacgtgtcc 720
cagatcgcca gctttcccgta tcagacagag gatggccggc tgccctagtc tggcaggatc 780
gtggtgact acatgatgca gaagcctggc aagacccggca ccatgtgttca ttagagggc 840
gtgctgtgc ctcagaaaagt gtggtgtgccc agcgccaggt ctaaagtgtat caaggccagc 900
ctgcctctga ttggcgaggc cgactgtctg cacgaaaagt acggccgcct gaacaagagc 960
aaggccctact acacaggcga gcacgccaag gccatcggttca attggcccat ctgggtgaaa 1020
accccccctga agctggccaa tggcacaag tacagacccctc cggccaaagct gctgaaagag 1080
agaggcttct ttggcgccat tgccggattt ctggaaaggcg gctgggaggg aatgattggcc 1140
ggctggcactg gctatacatc tcatggggcc catggcgtgg ctgtggccgc cgatctgaag 1200
tctacccagg aagccatcaa caagatcacc aagaacctga acagcctgag cgagctggaa 1260
gtgaagaatc tgcagagact gagcggccgc atggatgagc tgcacaacga gatctggaa 1320
ctggacgaga aagtggatga tctccgcgc gataacaatt cctcccgat tgaactggcc 1380
gtgctgtgtt ccaacgaggc catcatcaac agegaggatg aacacctgct ggcctggaa 1440
cggaagctga agaagatgct gggccctctt gccgtggaga tggcaacgg ctgcttcgag 1500
acaaaggcaca agtgcacca gacctgcctg gatagaatcg cggctggcacttcaatgcc 1560
ggcgagttca gcctgcctac ctgcacagc ctgaatatacc 1602

<210> SEQ ID NO 29
<211> LENGTH: 534
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 29

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
 35 40 45

-continued

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
 50 55 60
 Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
 65 70 75 80
 Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
 85 90 95
 Ser Ile Leu His Glu Val Lys Pro Val Thr Ser Gly Cys Phe Pro Ile
 100 105 110
 Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
 115 120 125
 Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
 130 135 140
 Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
 145 150 155 160
 Ala Thr Ser Lys Ser Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
 165 170 175
 Lys Asp Asn Asn Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
 180 185 190
 Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
 195 200 205
 Ser Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn Pro
 210 215 220
 Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
 225 230 235 240
 Gln Ile Gly Ser Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
 245 250 255
 Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
 260 265 270
 Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
 275 280 285
 Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
 290 295 300
 Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
 305 310 315 320
 Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
 325 330 335
 Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
 340 345 350
 Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala
 355 360 365
 Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly
 370 375 380
 Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys
 385 390 395 400
 Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu
 405 410 415
 Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp
 420 425 430
 Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu
 435 440 445
 Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser
 450 455 460

-continued

<210> SEQ ID NO 30

<211> LENGTH: 1602

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 30

ggtgatattt aggctgtcga aggtaggcag gctgaactcg cggcattga aggtgcacgc
ggcgattcta tccaggcagg tctgggtgca ctgtgtctt gtctcgaaagc agccgttgcc
120
gatctccacg gcagaagggc ccagcattt cttagtcgatc cgttccaggg ccagcagggt
180
ttcatactcg ctgttgatga tgccctcggt ggacaggcgc acggccagtt caatctggga
240
ggaaattgtt tcggcgccga gatcatccac ttctcgatcc agttccagga tctcgatgt
300
cagctcatcc atggcgccgc tcagtctcg cagattcttc acttcacgt cgctcaggct
360
gttcagggttc ttgggtatct tggttgatggc ttctgggtt gacttcagat cggcgccac
420
agccacgcca tgggccccat gagatgtata gccgtgccag cggcaatca ttccctccca
480
gccgccttc agaaatccgg caatggcgcc aaagaagcct ctctcttca gcagcttggc
540
gggagggtctg tacttggtgc cattggccag cttaggggg gtttccaccc agatggggca
600
attgcccatttgcgcgt gcctggcgatc gtatgtggc ttgtcttgc tcaggccggc
660
gtacttttcg tgcagacagt cgccctcgcc aatcagaggc aggctccct tgatcacatt
720
agacctgccc ctggcacacc acactttctg aggcagcagc acgcctctct gatacacat
780
ggtgccggtc ttgcagggtc tctgcatttgcat gtatgtccacc acgtatctgc cagactgagg
840
cagttccggca tcctctgtct gatcgaaa gctggccatc tgggacacgt agtgggtgg
900
cacggccatttgcgcgt taaacttctg ggggttgc tgcctgtaca gtttccat
960
ctgggtcttgc tgcgtcgatc ggaagccca cactgtgatc tgatctcgcc cctctgtaca
1020
gtatgtggc acctccacgg tcaggggtt ggtggccgtt ttgttgttgc ctttaggcac
1080
ggcccccaggcc attgtggcaa aaaagccgt ctggagggtg gcatgggac aagagccgt
1140
ggtgcccaatgt ctataagggtc cgccaggggc ttctcgatca tgcacatcat tctgggtgg
1200
cagccggatgt ttctcgatgc ctctcagcag gttggcagc tgcgtatct tggccggc
1260
gtgcgtatgt gggaaaggcagc cgctggtcac aggcttcact tgcgtgcagga tgctggc
1320
ggcagatggt gtggtgccca cacacatagg tctgcccaga gccacatcca gatcggtgc
1380
gttcaggccaa tcggggcaca gcttgcctt ggttctgggt cccttcagggt tggcgaagta
1440
gtcttgggttgc ggggtgggtgg tcagggat caccgttc acatttcactt cggccctgggt
1500
ggctgttttc accacatggg ggctattgtc gctggatgc cgggtgcaga ttctatcgcc
1560
qttqctqgtt accacatggg ggctattgtc gctggatgc cgggtgcaga ttctatcgcc
1602

<210> SEQ ID NO 31

-continued

<211> LENGTH: 1493
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 31

atgacaactc aacagccacg ctctgcttgg ggcaccatgc cgtccctaac gggaccattg	60
tgaaaaccat tactaacgtat cagatagagg tgactaatgc caccgagctg gtgcaaagta	120
gctccacagg agagatctgc gatagtcccc accagattct ggacggaaag aattgtacgc	180
tgatcgacgc gctgtgggc gaccctcagt gtgacggatt tcagaataag aagtgggatc	240
tgtttgtgga aaggtaaaag gcttattcaa attgctaccc ttacgtatgt cctgattatg	300
ccagcctgcg gtccctcgtc gcgtcttagt ggactctgga gttcaacaac gagtcattha	360
actggactgg cgttacacag aacggacta gttccgctt cataaggaga agcaaaaata	420
gtttcttcag cagactgaat tggctgacac atctgaactt caagtaccct gcactgaatg	480
taaccatgcc caacaacgag cagttcgata agctttacat ttggggagtt catcatctg	540
gcactgacaa ggatcagatc ttctgtat cccaggcttc cggcaggatt accgtgtcta	600
caaagagaag ccagcaaaact gtgtctccca atatcgccg tagaccaga gtacggaaca	660
tccctagtcg catcgtatt tactggacca tcgtgaaacc aggcgatatt ctcctgatta	720
acagtaactgg caacctgatc gcccccccgg gatactttaa aatccgctct ggaaagtccct	780
ccattatgag atcagatgca ccgatcgaa aatgcaactc tgagtgtatc acacccaatg	840
ggagcattcc caatgacaaa ctttccaga acgttaatcg aataacttat gggcctgtc	900
cacggtaactg gaagcaaaat accttgaaac tggcgaccgg tatgcgcaat gtcccccggaa	960
aacagacccg cggatattt gggctatcg caggctttat cgagaatggc tggaaaggaa	1020
tggtggatgg ttggatggg tttagacatc aaaactccga aggcagagcc caggctggcg	1080
atctcaagag cacgcaggcc gctatagatc agatcaatgg aaagctcaac agactgatcg	1140
ggaaaaaccaa cgaaaaattc catcagatcg agaaagatctt ctccgaagtc gagggcgca	1200
tacaggaccc ggagaagttt gttgaggata caaagattga tctgtggtcc tacaatgccc	1260
agctgttgtt ggctctggag aatcagcaca ctattgacct gaccgattca gagatgaaca	1320
aacttttga gaagacgaag aagcagctt aagaaatgc agaggacatg gggAACGGAT	1380
gctttaaat atatcataag tgtgataatg cctgcatacg atcaatttgc aatggatcc	1440
atgatcacga tgtttacagg gacgaagcgc tgaataacag gttccagata aaa	1493

<210> SEQ ID NO 32
<211> LENGTH: 519
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 32

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Ala			
1	5	10	15
Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly			
20	25	30	
His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp			
35	40	45	
Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr			
50	55	60	
Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Lys Asn Cys			
65	70	75	80

US 9,441,019 B2

119

120

-continued

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
 85 90 95

 Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
 100 105 110

 Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
 115 120 125

 Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
 130 135 140

 Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Lys
 145 150 155 160

 Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Asn Phe Lys
 165 170 175

 Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Gln Phe Asp Lys
 180 185 190

 Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Lys Asp Gln Ile
 195 200 205

 Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
 210 215 220

 Ser Gln Gln Thr Val Ser Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
 225 230 235 240

 Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
 245 250 255

 Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
 260 265 270

 Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
 275 280 285

 Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
 290 295 300

 Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
 305 310 315 320

 Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
 325 330 335

 Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala
 340 345 350

 Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
 355 360 365

 Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala Ala Asp Leu Lys
 370 375 380

 Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu
 385 390 395 400

 Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
 405 410 415

 Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
 420 425 430

 Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
 435 440 445

 Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
 450 455 460

 Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
 465 470 475 480

 Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
 485 490 495

 Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu

-continued

500

505

510

Asn Asn Arg Phe Gln Ile Lys
515

<210> SEQ ID NO 33
<211> LENGTH: 1493
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 33

ttttatctgg aacctgttat tcagcgcttc gtcctgttaa acatcgtgat cataggtacc	60
atttctaatt gatccgatgc aggccattatc acacttatga tatattttaa agcatccgtt	120
ccccatgtcc tctgcatttt ctctaagctg cttcttcgtc ttctcaaaaa gtttgttcat	180
ctctgaatcg gtcaggctaa tagtgtgctg attctccaga gccaccagca gctccggcatt	240
gtaggaccac agatcaatct ttgtatccctc aacatacttc tccaggtctt gtatgcgcc	300
ctcgacttcg gagaactctt tctcgatctg atgaaatttt tctgttggtt tcccgatctg	360
tctgttgagc tttccattga tctgatctat agccgcctgc gtgtcttga gatccggcagc	420
ctggccctcg ctttggagt tttgtatgtct aaaaccatac caaccatcca ccatcccttc	480
ccagccatcc tcgataaagc ctgcgcatacg cccaaatatac ccgcgggtct gttttcggg	540
gacattgcgc atacccggcg ccagttcaa ggtatttgc ttcacgtacc gtggacaggc	600
cccataagtt attcgattaa cgttctggaa aggtttgtca ttgggaatgc tcccatggg	660
tgtgatacac tcagagttgc atttccatcg cggatctatc gatctataa tggaggactt	720
tccagagccg atttaaagt atccccgggg ggcgcattcagg ttgcgcgtac tgttaatcag	780
gagaatatcg ctttgcgttca cgtatgttca gtaaatactg atgcgcactg ggtatgttccg	840
tactctgggt ctactgcgca tattgggaga cacagttgc tggcttctct ttgttagacac	900
ggtaatccctg ccggaaaggct gggcatacag aaagatctg tccctgtcag tgccaggatg	960
atgaactccc caaatgtaaa gcttatcgaa ctgctcggtt ttgggcattgg ttacattcag	1020
tgcagggtac ttgaagttca gatgtgtcag ccaattcagt ctgctgaaga aactatttt	1080
gtttctcctt atgcaagccg aactagccc gttctgttca acgccagttcc agttaaatga	1140
ctcggttgcg aactccagag tcccaactaga cgcgcacgagg gaccgcaggc tggcataatc	1200
aggcacatcg taagggttagc aatttgaata agcccttgac ctttccacaa acagatccca	1260
cttcttattc tggaaatccgt cacactgagg gtcgcacaa acgcgcgtca tcagcgtaca	1320
attcttcccg tccagaatct ggtggggact atgcgcagatc tctcctgtgg agctactttg	1380
caccagctcg gtggcattag tcacctctat ctgatcgtta gtaatgggtt tcacaatgg	1440
cccggttaggg acggcatggt gccccaaagca gagcgtggct gttgagttgt cat	1493

<210> SEQ ID NO 34
<211> LENGTH: 1551
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 34

atgaaaagtga agctgctggc gctgctgtt acctttaccg ccacccatcg cgcataccatc	60
tgtatcggtt accacgccaa caatagcacc gacaccgtgg ataccgtgct ggaaaagaac	120
gtgaccgtga cccacacgtt gaaacctgtg gaaaacagcc acaacggcaa gctgtgtcg	180
ctgaaaggca ttgcctctt gcagctggaa aattgtagcg tggccggctg gattctggc	240

-continued

aatcctgagt gcgagctgt gatttccaaa gagtccttgt cctacatcg ggagaagccc
aacccctgaga atggcacctg ctaccctggc cacttcgcg attacgagga actgagagaa
cagctgtccca gcgtgtccag cttcgagaga ttcgagatct tccccaaaga gagcagctgg
cccaatcata cagtgaccgg cgtagcgcc tctttagcc acaatggcga gagcagcttc
tacagaaacc tgctgtggct gacccggcaag aacggcctgt accccaacct gagcaagagc
tacgccaaca acaaagaaaa agaagtgtg gtctctggg gagtgcacca ccctcttaac
atcggcatcc agaaggccct gtaccacacc gagaatgcct acgtgtccgt ggtgtccagc
caactacgca gaaatgtcac ccccgagatc gccaaaagac ccaaagtgcg ggaccaggaa
ggcaggatca actactactg gaccctgctg gaacctggcg acaccatcat cttcgaggcc
aacggcaatc tgatcgcccc tagatacgcc tttgcccgtga cgagaggctt tggcgacggc
atcatcaaca gcaacgcccc catggacaag tgtgacgcca agtgtcagac accacaggga
gtctatcaata gcaagcctgcc cttccagaat gtgcacccctg tgaccatcg cgagtgtccct
aaatacgtgc ggagcgc当地 gctgagaatg gtgaccggcc tgaggaatat ccccagcatc
cagagcagag ggctgtttgg cgccattgcc ggctttatcg agggcggatg gacaggcatg
gtggatgggt ggtacggcta ccaccaccag aatgaggcagg gatctggcta tgccggcgat
cagaagagca cccagaacgc catcaacccgc atcacaaca aagtgaacag cgtgtatcg
aagatgaaca cccagttcac cgccgtggc aaagagttca acaagctgga acggcggatg
ggaaacctga acaagaaggt ggacgacggc ttcatcgaca tctggaccta caacgccc当地
ctcctggatcc tcctggaaaa tgagaggacc ctggacttcc acgacacgaa cgtgaagaac
ctgtacgaga aagtgaagag ccagctgaaag aacaacgc当地 aagagatcg当地
ttcgagttct accacaagtg caacgacgag tgcatggaaa gcgtgaagaa cggcacctac
gactaccata aagtacggca ggaaacggca agtgaacccgg agaaatcgatcgatcg
1551

<210> SEQ ID NO 35

<211> LENGTH: 517

<212> TYPE: PRT

<212> TYPE: PRI
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 35

Met Lys Val Lys Leu Leu Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1 5 10 15

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
 20 25 30

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
35							40					45			

Leu Leu Glu Asn Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile
50 55 60

Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly
65 70 75 80

Asn	Pro	Glu	Cys	Glu	Leu	Leu	Ile	Ser	Lys	Glu	Ser	Trp	Ser	Tyr	Ile
				85					90				95		

Val Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe
100 105 110

Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
 115 120 125

Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140

-continued

Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe
 145 150 155 160
 Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn
 165 170 175
 Leu Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu
 180 185 190
 Trp Gly Val His His Pro Pro Asn Ile Gly Ile Gln Lys Ala Leu Tyr
 195 200 205
 His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
 210 215 220
 Lys Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu
 225 230 235 240
 Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
 245 250 255
 Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala
 260 265 270
 Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met
 275 280 285
 Asp Lys Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300
 Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320
 Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335
 Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
 340 345 350
 Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365
 His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
 370 375 380
 Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
 385 390 395 400
 Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415
 Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile
 420 425 430
 Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
 435 440 445
 Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
 450 455 460
 Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
 465 470 475 480
 Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
 485 490 495
 Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
 500 505 510
 Arg Glu Lys Ile Asp
 515

<210> SEQ ID NO 36

<211> LENGTH: 1551

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 36

```

atcgatctc tcccggttca gcttgcttc ctcgctgtac ttggggtagt cgtaggtgcc      60
gttcttcacg cttccatgc actcgctgtt gcacttgtgg tagaactcga agcagccgtt     120
gcgcgcgtct ttggcggtgt tcttcagctg gctttcaact ttctcgtaa ggttttcac      180
gttgctgtcg tggaagtcca gggcccttc atttccagg aggaccagga gttcggcggt     240
gttaggtccag atgtcgatga agccgctcgac caccttcttg ttcagggttt ccattccgccc   300
ttccagcttg tgtaacttctt tgcccacggc ggtgaactgg gtgttcatct tctcgatcac   360
gtgttcaact ttgttggta tgccgtttagt ggccgtttagt gtgttctttt gateggccgc   420
ataggccatgc ccctgcgtcat tctgggtgg gtagccgtac caccatcca ccatgcctgt    480
ccatccgccc tcgataaaagc cggcaatggc gccaaacagg cctctgtctt ggatgttggg   540
gatattccctc aggccggtca ccattctcgat cttggcgttc cgacacgtt taggacactc   600
gcgcgcgttc acagggtgca cattctggaa gggcaggctg ctattgtatag ctccctgtgg   660
tgtctgacac ttggcggtcac acttgcctat gggggcggtt ctgtttagta tgccgtgcc   720
aaagccctcg otcagggcaa aggccgtatct aggggcgttcc agattggccgt tggccctcgaa 780
gtatgttggt tcgcccagggtt ccaggcagggtt ccagtagtag ttgtatctgc ctccctggtc 840
ccgcactttg ggtctttgg cgatctcggtt ggtgaacttt ctgtttagt ggctggacac   900
cacggacacg taggcattct cgggtggta cagggccctt tggatgcccgtt tgtaggagg   960
gtgggtcaact cccagagga ccagcacttc ttttctttt ttgttggccgt agctttgtct 1020
cagggtgggg tacaggccgt tcttgcgggtt cagccacacg aggtttctgtt agaagctgtt 1080
ctcgccattt tggctacaag aggegctcac gccggtaactt gtatgttggt gcaagctgtt 1140
ctctttgggg aagatctcgaa atctctcgaa gctggacacg ctggacacgtt gttctctcgat 1200
ttcctegtaa tcggcgaagt ggcagggttgcaggttgccttca tgggttttc 1260
cagcatgtt gaccaggact ctttggaaat cagcagctcg cactcaggat tgcccaagaat 1320
ccagccggcc acgctacaat ttcccagctg cagagggccatgccttca gcagacacacg 1380
cttgcgggtt tggctgtttt ccaggcagggtt cagcgttggt gtcacgggtca cggttttc 1440
cagcaeggtt tccacgggtt cgggtctatt gttggcgtgg tagccgatatac agatggatc 1500
ggcgttaggtt gcggtaaagg tacacagcag caccagcagc ttcaacttca t 1551

```

<210> SEQ ID NO 37

<211> LENGTH: 1605

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 37

```

atgaaggcca tcatcggtt gctgtatgggtt gtcacaaggca acggccgtatag aatctgtacc 60
ggcatcacca gcagcaatag ccctcacgtc gtggaaaacac ctacacaggcg cgaagtgtat 120
gtgaccggcg tgatccctctt gaccacaaca cctacaaaga gccacttcgc caatctgtat 180
ggcacagaga caagaggca gctgtgtccc aagtgcctga attgcacaga tctggatgtt 240
gtcttggca gacctaagtgtt tacaggcaaa atcccttagcgcc coagagtgttcc cattctgtat 300
gaagtgcgtt cttgtgttgc cgggtgtttt cctattatgc acgaccggac caagatcgat 360
cagctgccta atctgtgttcc aggttacggac cacatcgatgc tgagcaccatca caatgtgtat 420
aacggccaaa atgcttccgtt cggcccttat aagatcggttca catctggcagc ctggcccaac 480
attacaaatg gcaatggctt ctttgcaccatggcttggccgttccaa gaacgataag 540

```

US 9,441,019 B2

129

130

-continued

```

aacaagaccc ccaccaaccc cctgacaatc gaggtgccat atatctgtac agagggcgag 600
gatcagatca ccgtgtgggg atttcacagc gacaacgaaa cacagatggc caagctgtac 660
ggcgatagca agcctcagaa gtttaccagc tctgccaatg gcgtgaccac acactatgtg 720
tctcagatcg gcgggttccc taatcagaca gaagatggcg gactgectca gtctggaaga 780
atcgtggtgg attacatggt gcagaagtct ggcaagaccg gcaccatcac atatcagaga 840
ggaatcctgc tgccccagaa agtgtggcgc gtttctggaa gatccaaagt gatcaaggc 900
agcctgcctc tgattggaga agccgattgt ctgcacgaga aatacggcgg cctgaacaag 960
agcaaggcctt actatacagg cgagcacgccc aaggccatcg gcaattgtcc tatttggc 1020
aagacccttc tgaagctggc caatggcaca aagtatagac ctccagccaa gctgctgaaa 1080
gagagaggct tttttggc tatcgccgc tttctggaa gggatgggatgggatgatt 1140
gctggatggc atggctacac atctcatggc gcacatggcg tggcagtggc tgctgatctg 1200
aaatctacac aggaagccat caacaagatc accaagaacc tgaacagcct gagegagctg 1260
gaagtgaaga atctgcagag actgtctggc gccatggacg aactgcacaa tgagatcctg 1320
gaactggacg agaagggtggc cgatctgaga gcccatacaa tcagcagccaa gattgaactg 1380
gctgtgctgc tgtctaacga gggcatcatc aatagcgagg acgaacatct gctggccctg 1440
gaaagaaaagc tgaagaagat gctgggacct agcggcgtgg aaatcggcaa tggatgctt 1500
gagacaaagc acaagtgc当地 ccagacctgc ctggatagaa ttggccccc当地 aacatttgat 1560
gccggcgagt tttctctgcc caccttcgat agcctgaata tcaca 1605

```

<210> SEQ ID NO 38

<211> LENGTH: 535

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 38

Met	Lys	Ala	Ile	Ile	Val	Leu	Leu	Met	Val	Val	Thr	Ser	Asn	Ala	Asp
1															
					5				10					15	

Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys
					20				25					30	

Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Thr
					35			40			45				

Thr	Thr	Pro	Thr	Lys	Ser	His	Phe	Ala	Asn	Leu	Lys	Gly	Thr	Glu	Thr
					50			55			60				

Arg	Gly	Lys	Leu	Cys	Pro	Lys	Cys	Leu	Asn	Cys	Thr	Asp	Leu	Asp	Val
					65			70		75		80			

Ala	Leu	Gly	Arg	Pro	Lys	Cys	Thr	Gly	Lys	Ile	Pro	Ser	Ala	Arg	Val
					85			90			95				

Ser	Ile	Leu	His	Glu	Val	Arg	Pro	Val	Thr	Ser	Gly	Cys	Phe	Pro	Ile
					100			105			110				

Met	His	Asp	Arg	Thr	Lys	Ile	Arg	Gln	Leu	Pro	Asn	Leu	Arg	Gly
					115			120			125			

Tyr	Glu	His	Ile	Arg	Leu	Ser	Thr	His	Asn	Val	Ile	Asn	Ala	Glu	Asn
					130			135			140				

Ala	Pro	Gly	Gly	Pro	Tyr	Lys	Ile	Gly	Thr	Ser	Gly	Ser	Cys	Pro	Asn
					145			150		155			160		

Ile	Thr	Asn	Gly	Asn	Gly	Phe	Phe	Ala	Thr	Met	Ala	Trp	Ala	Val	Pro
					165			170			175				

Lys	Asn	Asp	Lys	Asn	Lys	Thr	Ala	Thr	Asn	Pro	Leu	Thr	Ile	Glu	Val
					180			185			190				

-continued

Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe
195 200 205

His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser Lys
210 215 220

Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val
225 230 235 240

Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro
245 250 255

Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys
260 265 270

Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val
275 280 285

Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu
290 295 300

Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys
305 310 315 320

Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys
325 330 335

Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr
340 345 350

Arg Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile
355 360 365

Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His
370 375 380

Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu
385 390 395 400

Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser
405 410 415

Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met
420 425 430

Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp
435 440 445

Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu
450 455 460

Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu
465 470 475 480

Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly
485 490 495

Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp
500 505 510

Arg Ile Ala Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro Thr
515 520 525

Phe Asp Ser Leu Asn Ile Thr
530 535

<210> SEQ ID NO 39

<211> LENGTH: 1605

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 39

tgtatattc aggctatcga aggtggcgag	agaaaaactcg ccggcatcaa atgttccggc	60
ggcaattcta tccaggcagg tctggttgca	cttgtgtttt gtctcaaagc atccattgcc	120

-continued

gatttccacg	gcgcttaggtc	ccagcattt	cttcagctt	ctttccaggg	ccagcagatg	180		
ttcgtctcg	ctattgtga	tgccctcggt	agacagcgc	acagccagg	caatctggct	240		
gctgattgt	ta	tcggtctca	gategtccac	cttctcg	agttccagg	300		
cagt	tcgtcc	atggggccag	acagtctcg	cagatttctc	acttccagct	360		
gttcaggttc	ttgggtatct	tgtgtatggc	ttcctgtgt	gatttcagat	cagcagcac	420		
tgccaa	cgccat	gagatgtgt	gecatgecat	ccagcaatca	ttccctcca	480		
tccgccttcc	agaaagccgg	cgatagctcc	aaaaaaagct	ctctcttca	gcagcttggc	540		
ttggagg	tcta	tacttgtgc	cattggccag	cttcagagg	gtcttacccc	600		
attggccat	gccttggcgt	gctcgctgt	atagtaagg	ttgtcttgt	tcaggccgc	660		
gtatttctcg	tgca	gacaat	cggcttctcc	aatcagagg	aggctgcct	720		
ggatcttcca	gaagcgcacc	acacttctg	gggcagcagg	attcctct	gatatgtgt	780		
ggtgccgg	tc	tgccagact	tctgcaccat	gtaatccacc	acgattctc	840		
cagtccgcca	tcttctgt	gattaggaa	gccgccc	tgagacacat	agtgtgtgg	900		
cacggccatt	gca	gagactgg	taaacttctg	aggcttgct	tcgccc	960		
ctgtgtttcg	ttgtcgctgt	gaaatcccc	cacgggtatc	tgatctc	cctctgtaca	1020		
gatata	atggc	acctcgatt	tcagggggtt	ggtggcggtc	ttgttcttat	cgttcttag	1080	
cacggcccaa	gccc	atgggtgt	caaagaagcc	attgcccatt	gtaatgttgg	ggcagctg	1140	
agatgtgc	atcttataag	ggccgc	aggat	ttcg	gcgttgc	cattgtgg	1200	
gctcagtctg	atgtgctcg	agcctctc	agat	ttcg	agctgtctg	tcttgg	1260	
gtcgtgcata	ataggaaa	ac	ggcgctgg	ca	aggtgc	acttcatg	1320	
tctggcgct	gggat	ttgc	tgt	ctgt	atgg	gacac	1380	
gcaattcagg	cacttgg	acag	cttgc	tctgt	gtgc	ccatt	1440	
gtggctt	gtaggtgtt	tgg	tgcagg	gtc	acatt	ctc	1500	
tgtagctgtt	ttcacgacgt	gagg	gtt	gtgtt	atgcgg	tac	1560	
ggcg	ttgt	gtgaccacca	tcagcag	ac	gatgatgg	cc	ttcat	1605

<210> SEQ ID NO 40

<211> LENGTH: 2058

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 40

atgaaggcca	aactgttgt	gctgttgt	ac	tttac	ccac	ctacgc	cgacacaatc	60		
tgtatcg	gct	accacg	ccaa	caat	agcacc	gacaccgt	tttt	ggagaagaac	120	
gtgaccgt	gta	ccc	actctgt	gt	aa	ccat	ggc	aa	gctgtgt	180
ctgaaagg	ca	tgc	ccc	ct	gc	agctgg	tttt	ggat	ctgg	240
aaccccg	gt	gag	ctgt	gt	ttt	ctaa	gag	ctgg	atgc	300
aatcctg	aga	at	ggcac	ct	acc	ctgg	tact	tc	gatc	360
cagctgt	cta	gc	gtgtcc	ag	ttc	cgag	ttt	ccaa	gtcc	420
ccta	atc	ca	cgac	agg	cg	tcgt	tc	at	ccat	480
taccgg	aa	cc	tg	tg	tt	gg	at	gg	tt	540
tacgtga	aca	aa	gg	aa	gt	gt	gg	tt	gg	600

-continued

atcggaaatc agccccccct gtaccacaca gagaacgcgt atgtgagcgt ggtgtccagc
caactacagca gaagattcac cccc gagatc gccaagagac ccaaagttag agaccaggag
ggccggatca attactactg gaccctgctg gagcttgcg ataccatcat ctgcaggcc
aacggcaatc tgatcgcccc ttggtatgcc tttgcctgaa gcagaggctt tggcagcggc
atcatcacaa gcaacgcccc catggatgag tgtgatgcca agtgcacagac acctcaggcc
gccccatcaata gcagcctgcc cttccagaat gtgcacccctg tgaccatcg cgagtgc
aagtatgtga gaagcgccaa gctgagaatg gtgacccggc tgagaaacat ccctcaggagg
gagaccagag gactgtttgg agccatcgcc ggattcatcg agggaggatg gacaggcatg
gtggatggct ggtacggcta ccaccaccag aatgagcagg gctctggata tggcccgat
cagaagtcta cccagaacgc catcaacggc atcaccaaca aggtgaacag cgtgatcgag
aagatgaaca cccagttac cgctgtggc aaggagttca acaagctgga gcgaggatg
gagaacctga acaagaaggt ggacgacggc tttctggaca tctggaccta caatccgaa
ctcctggtcc tcctcgagaa tgagaggacc ctggacttcc acgacacgaa cgtgaagaac
ctgtatgaga aggtgaagag ccagctgaa aacaacgcca aggagatcg caacggctgc
ttcgagttct accacaagtg taacaacgag tgtatggaga gctgtaaagaa cggcacctac
gactacccta agtacagcga ggagagcaag ctgaaccggg agaagatcga ttccggaggc
gacatcatca agctgtgaa cgagcagggt aacaaggaga tgcagagcag caaccgtac
atgagcatga gcagctggc ctacacccac agectggacg ggcggccct gttctgttc
gaccacgccc cggaggagta cgagcacgac aagaagctga tcatcttctt gaaaggaaac
aacgtgccc tgca gtcagctgac cagcatcagc gccccggc acaagttcga gggcttgacc
cagatcttcc agaaggccctt cgagcacgag cagcacatca gcgagagcat caacaacatc
gtggaccacg ccatcaagag caaggaccac gccaccccttca acttcttgca gtggatcg
ggcgagcgc acgaggaggaa ggtgtgttc aaggacatcc tggacaagat cgagctgatc
ggcaacgaga accacggccctt gtacccggcc gaccagtagc tgaaggccat cgccaaagagc
aaqaaqaqccq qatcctaaq

```
<210> SEQ ID NO 41
<211> LENGTH: 685
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 41

1 5 10 15
Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asp Asn Ser Thr Asp Thr

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn

Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile

Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly

Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile
25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40

-continued

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100 105 110

Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140

Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn
165 170 175

Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu
180 185 190

Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr
195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
210 215 220

Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu
225 230 235 240

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
245 250 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala
260 265 270

Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
275 280 285

Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
290 295 300

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
305 310 315 320

Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
325 330 335

Ile Pro Gln Arg Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
340 345 350

Ile Glu Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
370 375 380

Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
385 390 395 400

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
420 425 430

Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
435 440 445

Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
450 455 460

Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
465 470 475 480

Phe Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

Arg Glu Lys Ile Asp Ser Gly Gly Asp Ile Ile Lys Leu Leu Asn Glu

US 9,441,019 B2

139

140

-continued

515	520	525
Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met Ser		
530	535	540
Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe		
545	550	555
Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe		
565	570	575
Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro		
580	585	590
Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu		
595	600	605
His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala		
610	615	620
Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val		
625	630	635
Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys		
645	650	655
Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln		
660	665	670
Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser		
675	680	685

<210> SEQ ID NO 42
 <211> LENGTH: 2058
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 42

ctaggatccg ctcttcctgc tcttggcgat gcccgttacag tactggtcgg ccaggtacag	60
gccgtggttc tcgttgccga tcagctcgat cttgtccagg atgtccttga acagcacctc	120
ctcctcggtgc tgctcgccca cgttaccactg caggaagtgg aagggtggcgt ggtccttgc	180
cttggatggcg tggtccacga tggtgttcat gctctcgctg atgtgtgtct cgtgtcgta	240
ggcccttctgg aagatctggg tcaggccctc gaacttgtgc tggggggcgc tgatgtgg	300
cagctgcacg ggcacgttgc tctcggtcag gaagatgtac agtttttgg cgtgtcgta	360
ctcctcggtcg acaggaacag gcccggcccg tccaggctgtt ggggttagca	420
ccagctgctc atgtcatgt acagggtgt gctctgcattt tccctgttca cctgtcgctt	480
cagcagcttgc atgtatgtcgc ctccggaaatc gatcttctcc cggttcagct tgctctcctc	540
gtgttactta gggtagtgcgtt aggtggccgtt ctgcacgttgc tccataactt cgttgttaca	600
cttggtagtgcgtt aactcgaagc agccgttgc gatcttgcgtt ggtttgttct tcaagttgcgtt	660
cttcacatttc tcatacaggat tcttcacgtt gctgtcggtt gaaatccagg tccatcttca	720
ctcgaggagg accaggaggat cggcattgtt ggtccagatg tccagaaaagc cgtcgacac	780
cttcttgcgttcc aggttgcgttcc tccctccgtc cagttgttgc aactccttgc ccacagcggt	840
aaactgggttgc ttcatcttctt cgttacgttgc gttcacgttgc ttgggtatgc cgttgcgttgc	900
gttctgggttgc gacttctgtat cggccggata tccagagccc tgctcattttt ggtgggttgc	960
ccgttaccatc ccatccacca tgcctgttca tccctccgtc atgaatccgg cgttgcgttgc	1020
aaacagtccttcttctt cttggctccc tctggggat gtttctcagg cgggtcacca ttctcagtt	1080
ggcgcttctc acataacttgg ggcactcgcc gatggtcaca gggtgcacat tctggaaagg	1140

-continued

caggctgcta ttgtatggcgc cctgagggtgt ctggcacattt gcatcacact catccatggg	1200
ggcggttgctt gtgtatgtgc cgctgccaaa gcctctgtc agggcaaagg cataccaagg	1260
ggcgatcaga ttgcggttgg cctcgaagat gatggtatcg ccaggctcca gcagggtcca	1320
gtagtaattt atccggccct cctggtctct cactttgggt ctcttggcga tctcggggt	1380
aatcttcgt ctgttagtggc tggacaccac gtcacatag gcgttctctg tgtggtacag	1440
ggcccgctga ttcccgatgt tggggggtg gtgcactccc cacagcacca gcactccctt	1500
ttccttgttg ttcacgttagc tcttgcttag gttgggttac aggccattct tgcctgtcag	1560
ccacagcagg ttcccgtaga agctgtttt gccgttgtgg ctacagctgg cagacacgccc	1620
tgtcactgtg tgattaggcc agctggactc ctggggaaag atctcaatc tctcaagct	1680
ggacacgctta gagacgtgtc cgccgagtc ctcgtaatcg gcaagatgc cagggttagca	1740
ggtgccattc tcaggatggg gggtctccac gatgttagtc cagctctct tagaatcag	1800
cagctcacac tcggggttgc ccagaatcca tccggccaca gaacaattgc ccagctgcag	1860
agggggcaatg ccttcagca gacacagctt gccattgtgg ctgtcttcca gcaggttcac	1920
agagtgggtc acggtcacgt tcttctccag cactgtatcc acgggtgtcg tgctattgtt	1980
ggcgtggtag ccgatacaga ttgtgtcggc gtaggtggcg gtaaaaggta acagcagcac	2040
caccaatggggccat	2058

```
<210> SEQ ID NO 43
<211> LENGTH: 2061
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 43

atgaaggcca tcctgggttgt gctgtgtac acttcgcca cggccaaacgc cgacaccctg
tgcatecggtt accacgccaa caacagcacc gacaccgtgg acaccgtgt ggagaagaac
gtgaccgtga cccacagcgt gaacctgtct gaggacaage acaacggcaa gctgtcaag
ctgcggggcg tggccccctt gcacctgggc aagtgcaca tcgcccgtg gattctggc
aaccggagt gcgagagcct gagcaccggc agcagctggta gtcatacatgt ggagaccccc
agcagcggaca acggcacctg ctacccggc gacttcatcg actacgagga gctgcccc
cagctgagca gctgtggcag cttegagcgg ttggatgtat tcccaagac cagcagctgg
cccaaccacg acagcaacaa gggcgtgacc gcccgtgcc cccacggcgg cgccaagagc
ttctacaaga acctgtatctg gctggtgaaag aaggcaaca gtcacccaa gctgagcaag
agtcatacatca acgacaaggg caaggagggtg ctgggtgtgt ggggcatcca ccaccc
accagcgcgg accagcagag cctgtaccag aacgcccaca cctacgttt cgtggcagc
agccggtaca gcaagaagtt caagcccgag atcgccatcc ggcccaaggt ggggaccag
gagggccggta tgaactacta ctggaccctg gtggagcccg gegacaagat caccc
ggccacccggca acctgggttgt gccccgggtac gcctcgcca tggagccgaa cggccggc
ggcatcatca tcagcgacac ccccggtcaca gactgcaca ccacccgtcc gaccccaag
ggccatca acaccaggcct gcccgtcccg aacatccacc ccatcaccat cggcaagtgc
cccaagtacg tgaagagcac caagctgcgg ctggccacccg gctgccccggaa catcccc
atccagagcc gggccgttt cgccgtccatc gcccgttca tcgagggccgg ctggaccggc
1020
1030
1040
1050
1060
1070
1080

-continued

atggtgacg gctgtacgg ctaccaccac cagaacgacg agggcagccg ctacccgc	1140
gacctgaaga gcacccagaa cgccatcgac gagatcacca acaaggtaa cagcgtatc	1200
gagaagatga acacccagtt caccggcgtg ggcaaggagt tcaaccacct ggagaagccg	1260
atcgagaacc tgaacaagaa ggtggacgc ggcttcctgg acatctggac ctacaacgcc	1320
gagctgtgg tgctgtgga gaacgagcc accctggact accacgacag caacgtaa	1380
aacctgtacg agaagggtgcg gagccagctg aagaacaacg ccaaggagat cggcaacggc	1440
tgttgcagt tctaccacaa gtgcgacaac acctgcattgg agagcgtgaa gaacggcacc	1500
tacgactacc ccaagtacag cgaggaggcc aagctaaacc gggaggagat cgactccgga	1560
ggcgacatca tcaagtgct gaacgagcc gtgaacaagg agatcagag cagcaacctg	1620
tacatgagca tgagcagctg gtgtacacc cacagcctgg acggcgccgg cctttcctg	1680
tgcaccacg ccggcggagga gtacgagcc gccaagaagg tgcatactt cctgaacgag	1740
aacaacgtgc ccgtgcagct gaccagcatc agegcccccg agcacaagtt cgagggctg	1800
acccagatct tccagaaggc ctacgagcc gacgacacatc tgcgcgagag catcaacaac	1860
atcggtggacc acgccccatcaa gagcaaggac cacgcccaccc tcaacttcct gcagttgtac	1920
gtggccgagc agcagcggagga ggagggtgtt tcaaggaca tccctggacaa gatcgagctg	1980
atcggtggacc acgccccatcaa gagcaaggac cacgcccaccc tcaacttcct gcagttgtac	2040
aqcaqqaaqa qcqqtatcta q	2061

```
<210> SEQ ID NO 44
<211> LENGTH: 686
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 44

Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Ala Thr Ala Asn
1 5 10 15

Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
20 25 30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
35 40 45

Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu Arg Gly Val
50 55 60

Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp Ile Leu Gly
65 70 75 80

Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp Ser Tyr Ile
 85 90 95

Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro Gly Asp Phe
 100 105 110

Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro Asn His Asp

Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly Ala Lys Ser

Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro

Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val

-continued

Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu
 195 200 205
 Tyr Gln Asn Ala Asp Thr Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser
 210 215 220
 Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln
 225 230 235 240
 Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys
 245 250 255
 Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe
 260 265 270
 Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro
 275 280 285
 Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn
 290 295 300
 Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys
 305 310 315 320
 Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg
 325 330 335
 Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
 340 345 350
 Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
 355 360 365
 His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser
 370 375 380
 Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile
 385 390 395 400
 Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His
 405 410 415
 Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe
 420 425 430
 Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
 435 440 445
 Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu
 450 455 460
 Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly
 465 470 475 480
 Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val
 485 490 495
 Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu
 500 505 510
 Asn Arg Glu Ile Asp Ser Gly Asp Ile Ile Lys Leu Leu Asn
 515 520 525
 Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met
 530 535 540
 Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu
 545 550 555 560
 Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile
 565 570 575
 Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala
 580 585 590
 Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr
 595 600 605

-continued

Glu	His	Glu	Gln	His	Ile	Ser	Glu	Ser	Ile	Asn	Asn	Ile	Val	Asp	His
610							615						620		
Ala	Ile	Lys	Ser	Lys	Asp	His	Ala	Thr	Phe	Asn	Phe	Leu	Gln	Trp	Tyr
625							630					635			640
Val	Ala	Glu	Gln	His	Glu	Glu	Glu	Val	Leu	Phe	Lys	Asp	Ile	Leu	Asp
								645				650			655
Lys	Ile	Glu	Leu	Ile	Gly	Asn	Glu	Asn	His	Gly	Leu	Tyr	Leu	Ala	Asp
								660			665			670	
Gln	Tyr	Val	Lys	Gly	Ile	Ala	Lys	Ser	Arg	Lys	Ser	Gly	Ser		
								675			680			685	

```
<210> SEQ ID NO 45
<211> LENGTH: 2061
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 45

ctaggatccg ctcttccctgc tcttggcgat gcccttacag tactggtcgg ccaggtacag
ggcgtggttc tcgttgccga tcagctcgat cttgtccagg atgtccttga acagcaccc
ctcctcggtgc tgctcgccca cgtaccactg caggaagtgg aagggtggcgt ggtcccttgc
cttgcgtggcg tggccacga tggttgtat gctctcgctg atgtgtgtct cgtgtcgta
ggcccttgg aagatctggg tcaggccctc gaacttgc tccggggcgc tgatgttgt
cagctgcacg ggcacgttgt tctcggttag gaagatgtac agcttcttgg cgtgtcgta
ctcctegggc gcgtggcgtca acaggaacag gecggccgc tccaggctgt gggtgtagea
ccagctgctc atgctcatgt acagggtgtct gctctgcata tccttgcata cctgtcggt
cagcagcttgc atgatgtcg ctcggagtc gatctctcc cggttcaagct tggtcttgc
gtgtacttgc gggtagtgcgt aggtggcgtt cttcacgttc tccatgcagg tggtgtcgca
cttgcgttag aactcgaagc agccgttgc gatctcttgc gctgtttct tcagtcggct
ccgcacccccc tcgtacagggt tcttgcacgtt gctgtcggtt tagtccaggg tccgtcggt
ctccagcagc accagcaggt cggcggttgc ggtccagatg tccaggaagc cgtgtccac
cttgcgttgc aggtctcg tccgcgttgc cagggtgggtt aactcttgc ccacggccgt
gaactgggtt ttcatcttct cgtacacgtt gttcaccttgc ttgggtatct cgtcgatggc
gttgcgttgc ctcttcagggt cggcggttgc gccgtgttgc tgcgtttct ggtgggttgc
gccgttccaccc cggccgttgc gccgttgcgttgc atgaagccgg cgtatggcc
gaacaggcccc cggctctggc tgctggggat gttccgcagg cccgtggcca gccgcagctt
ggtgcttcc acgtacttgg ggcacttgc gatgggtatgc ggggtggatgt tctggaaagg
caggctgggttgc ttgtatggcgc cttgggggttgc cttggcagggttgc gtgttgcgtt cgtgcaccc
gggtgtcgcttgc atgatgtatgc cgtgtccggc gttccgttgc atggcgaagg cgtaccgggg
caccaccagg ttgcgggtgg ccttgcgttgc gatcttgcgttgc cccgggttgc ccagggttgc
gtatgttttc atccggccctt ccttgcgttgc caccgttgcgttgc cggatggcgttgc tcttgggttgc
gaactcttgcgttgc tgcttgcgttgc cttgggggttgc cttggcagggttgc gtgttgcgtt cgtgcaccc
gctctgttgcgttgc tgcttgcgttgc cttgggggttgc cttggcagggttgc gtgttgcgtt cgtgcaccc
gcccttgcgttgc ttgtatgttgc tcttgcgttgc cttgggggttgc cttggcagggttgc gtgttgcgtt cgtgcaccc
ccagatcagggttgcgttgc cttgggggttgc cttggcagggttgc cttggcagggttgc gtgttgcgtt cgtgcaccc

-continued

cttggcgtg tcgtgggtgg gccagctgtc ggtcttgggg aagatctcgaa accgctcgaa	1680
gctgctcacg ctgctcagct gctcccgacg ctccctcgtag tgcgtgaagt cgccgggta	1740
gcagggtgccg ttgtcgctgc tgggggtctc cacgtatgtac ctccagctgc tggcggtgct	1800
caggctctcg cactcggtt tgcccagaat ccagccggcg atgttgact tgcccaggta	1860
caggggggcc acgccccgca gcttgcacag cttgcccgtt tgcttgtctt ccagcagggtt	1920
cacgctgtgg gtcacggtca cgttcttc cagcacgggtt tccacgggtgt cggtgtgtt	1980
gttggcgtgg tagccgatgc acagggtgtc gggttggcg gtggcgaagg tgtacagcag	2040
caccaccagg atggccttca t	2061

<210> SEQ ID NO 46
<211> LENGTH: 2049
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 46	
atggccatca tctacactgat cctgtgttt acagctgtgc ggggcgatca gatctgtatc	60
ggctaccacg ccaacaatag caccgagaag gtggacacca tcctggaaag aaatgtgacc	120
gtgaccacacg ccaaggatata tctggaaaag acccacaacg gcaagctgtg caagctgaat	180
ggcatttcctc ctctggaaact gggcgattgt tctattgtgc gctggctgct gggaaatcct	240
gagtgcgata gactgtgtc tggcctgag tggagctaca tcatggaaaa agagaaccct	300
agggacggac tgggttaccc cggcagcttc aacgattacg aggaactgaa gcacctgctg	360
tccagcgtga agcacttcga gaaagtgaag atcctgccc aggatagatg gaccacat	420
acaacaacag gcgaaacgag agcttgcgtc gtgtccggca accccagctt cttcagaat	480
atggtctggc tgaccaagaa gggctctaata tattctgtgg ccaagggcag ctacaataat	540
acaagcggcg agcagatgtt gattattgg ggcgtgcacc accctaataa tgagacagag	600
cagagaaccc tggaccatca tggggcaca tacgtgtctg tggcaccag cacactgaat	660
aagagaagca ccccgatata tgccaccaga cccaaatgtca atggacaggg cggcagaat	720
gaattttctt ggaccctgtc ggatatgtgg gacaccatca actttgagag caccggaaat	780
ctgattgccc ctgagttacgg cttcaagatc agcaagagag gcagcagcgg catcatgaaa	840
acagaggcga ccctggaaaa ctgtgaaacc aagtgtcaga cacctctgg cgccattaat	900
accacccctgc cttccataa tggcaccct ctgacaatcg gcgagtgc taagtacgt	960
aagtctgaga aactgggtct ggccacagga ctgagaaatg tgccccagat cgagtcaaga	1020
ggcctgtttt gaggcattgc cggctttattt gaaggcggat ggcaggaaat ggtggatgg	1080
tggtaacggct atcaccacag caatgatcag ggatctggct atgcccggca taaagagagc	1140
acccagaagg ctttgacgg catccaacaa aagtgtaca gctgtatcga gaatgtac	1200
acccagtttggc aggccgtggg caaagagttc agcaatctgg aaagacggct ggaaacactg	1260
aacaagaaaa tggaaatgtt cttctggac gtgtggacat ataatgccga gctgtgtgt	1320
ctgatggaaa acgagaggac cctggacttt cacgacacgca acgtgtacgac	1380
aaagtgcgga tgcagctgag agacaatgtt aaagagctgg gcaacggctg ctttgagttc	1440
taccacaatgtt ggcacacgca gtgtcgatgtt agcgtgtacgaa acggcaccata cgactaccc	1500
aagtatgagg aagagagcaa gctgaacaga aacgagatca agtccggagg cgacatcatc	1560
aagctgtgtca acgacggatgtt gaacaaggag atgcagacgtca gcaacctgtt catgacat	1620

-continued

```

agcagctggt gctacaccca cagcctggac ggccgcggcc tggcctgtt cgaccacgcc 1680
gccgaggagt acgagcacgc caagaagctg atcatcttcc tgaacgagaa caacgtgccc 1740
gtgcagctga ccagcatcg cggccccgag cacaagttcg agggcctgac ccagatctc 1800
cagaaggct acgagcacga gcagcacatc agcgagagca tcaacaacat cgtggaccac 1860
gccatcaaga gcaaggacca cgccaccccttc aacttcctgc agtggctacgt ggccgagcag 1920
cagcaggagg aggtgctgtt caaggacatc ctggacaaga tcgagctgat cggcaacgag 1980
aaccacggcc tgcacctggc cgaccagtagt gtgaaggcgtc tgccaaagag caggaagagc 2040
ggatcctag 2049

```

```

<210> SEQ ID NO 47
<211> LENGTH: 682
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```
<400> SEQUENCE: 47
```

```

Met Ala Ile Ile Tyr Leu Ile Leu Leu Phe Thr Ala Val Arg Gly Asp
1 5 10 15

```

```

Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Lys Val Asp
20 25 30

```

```

Thr Ile Leu Glu Arg Asn Val Thr Val Thr His Ala Lys Asp Ile Leu
35 40 45

```

```

Glu Lys Thr His Asn Gly Lys Leu Cys Lys Leu Asn Gly Ile Pro Pro
50 55 60

```

```

Leu Glu Leu Gly Asp Cys Ser Ile Ala Gly Trp Leu Leu Gly Asn Pro
65 70 75 80

```

```

Glu Cys Asp Arg Leu Leu Ser Val Pro Glu Trp Ser Tyr Ile Met Glu
85 90 95

```

```

Lys Glu Asn Pro Arg Asp Gly Leu Cys Tyr Pro Gly Ser Phe Asn Asp
100 105 110

```

```

Tyr Glu Glu Leu Lys His Leu Leu Ser Ser Val Lys His Phe Glu Lys
115 120 125

```

```

Val Lys Ile Leu Pro Lys Asp Arg Trp Thr Gln His Thr Thr Thr Gly
130 135 140

```

```

Gly Ser Arg Ala Cys Ala Val Ser Gly Asn Pro Ser Phe Phe Arg Asn
145 150 155 160

```

```

Met Val Trp Leu Thr Lys Lys Gly Ser Asn Tyr Pro Val Ala Lys Gly
165 170 175

```

```

Ser Tyr Asn Asn Thr Ser Gly Glu Gln Met Leu Ile Ile Trp Gly Val
180 185 190

```

```

His His Pro Asn Asp Glu Thr Glu Gln Arg Thr Leu Tyr Gln Asn Val
195 200 205

```

```

Gly Thr Tyr Val Ser Val Gly Thr Ser Thr Leu Asn Lys Arg Ser Thr
210 215 220

```

```

Pro Asp Ile Ala Thr Arg Pro Lys Val Asn Gly Gln Gly Arg Met
225 230 235 240

```

```

Glu Phe Ser Trp Thr Leu Leu Asp Met Trp Asp Thr Ile Asn Phe Glu
245 250 255

```

```

Ser Thr Gly Asn Leu Ile Ala Pro Glu Tyr Gly Phe Lys Ile Ser Lys
260 265 270

```

```

Arg Gly Ser Ser Gly Ile Met Lys Thr Glu Gly Thr Leu Glu Asn Cys

```

US 9,441,019 B2

153**154**

-continued

275 280 285

Glu Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Thr Thr Leu Pro
 290 295 300

Phe His Asn Val His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr Val
 305 310 315 320

Lys Ser Glu Lys Leu Val Leu Ala Thr Gly Leu Arg Asn Val Pro Gln
 325 330 335

Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
 340 345 350

Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Ser Asn
 355 360 365

Asp Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala
 370 375 380

Phe Asp Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn
 385 390 395 400

Thr Gln Phe Glu Ala Val Gly Lys Glu Phe Ser Asn Leu Glu Arg Arg
 405 410 415

Leu Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp Val Trp
 420 425 430

Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr Leu
 435 440 445

Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Met
 450 455 460

Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe
 465 470 475 480

Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly Thr
 485 490 495

Tyr Asp Tyr Pro Lys Tyr Glu Glu Ser Lys Leu Asn Arg Asn Glu
 500 505 510

Ile Lys Ser Gly Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn
 515 520 525

Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys
 530 535 540

Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala
 545 550 555 560

Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu
 565 570 575

Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys
 580 585 590

Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln
 595 600 605

His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser
 610 615 620

Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln
 625 630 635 640

His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu
 645 650 655

Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys
 660 665 670

Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 675 680

-continued

```

<211> LENGTH: 2049
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 48

ctaggatccg ctcttcgtc tcttggcgat gcccttcacg tactggtcgg ccaggtacag      60
gccgtggttc tcgttgcgcga tcgactcgat cttgtccagg atgtccttga acageacactc    120
ctcctcgtgc tgctcggcca cgtaccactg caggaagtgg aaggtggcggt ggtccttgc      180
cttggatggcg tggatggcgta tggtgttgcgt gctctcgctg atgtgtgtgt cgtgtcgta    240
ggccttcgtgg aagatctggg tcaggccctc gaacttgtgc tcgggggcgc tgatgtgggt    300
cagctgcacg ggcacgttgc tctcggtcag gaagatgtac agcttctgg cgtgtcgta      360
ctcctcggcg gcgtggtcga acaggaacag gcccggccgc tccaggctgt gggtagca       420
ccagctgcgtc atgctcatgt acagggtgtcgt gctctgcata tccttgcgtca cctgtcggt    480
cagcagtttgc atgatgtcgat ctcggactt gatctcgat cttgtcgatgt tgctcttc      540
ctcataactta gggtagtcgt aggtggcgat cttcacgtca ttcatgcact cgtagtcgca    600
cttgggttag aactcaaagc agccgttgcg cagcttttc acattgtcgtc tcaagtcgtc      660
ccgcactttgc tcgtacaggt tcctcacgtt gctgtcgta aagtccagg ggcttcgtt      720
ttccatcagc accagcagct cggcattata tggccacacg tccaggaaac catctccat    780
tttcttgcggtag aggttttcca gccgtcttc cagattgtcg aactcttc ccacggcctc    840
aaactgggtg ttcatcttc cgtacacgt gttcacatttgg tgggtatgc cgtcaaggc      900
cttctgggtg ctctcttat cggcggcata gccagatccc tgatcattgc tgggtgata      960
gccgtaccac ccattccacca ttccctgcca tccgccttca ataaaggccgg caatggctcc 1020
aaacaggcct cttgactcgat tctggggcac atttctcgt cctgtggccca gcaccagttt 1080
ctcagacttc acgtacttag ggcactcgcc gattgtcaga gggtgcacat tatggaaagg 1140
cagggtggta ttaatggcgc ccagagggtgt ctgacacttg gttcacagt tttccagggt 1200
gccctctgtt ttcatgtgc cgtgtcgcc tctcttgcgt atcttgcgtc cgtactcgagg 1260
ggcaatcaga ttccgggtgc tctcaaagtt gatgggtcgtc cacatatcca gcagggtcca 1320
ggaaaattcc attctgcgc cctgtccatt cactttgggt ctgggtggcaa tatcgggggt 1380
gtttctctta ttcaagtgtgc tgggtggccac agacacgtat gtgcccacat tctggtacag 1440
ggttctctgc tctgtctcat cattagggtgt gtgcacgcggc caaataatca gcatctgtc 1500
gecgcttgta ttattgttagc tgcccttggc cacaggataa ttagagccct tcttggtgc 1560
ccagaccata ttctgaaga agctgggggtt gcccggacaca gcacaagcgtc tcttccggcc 1620
tgggtttgtt tgcgtgggtcc atctatcctt gggcaggatc ttcaacttgc cgaagtgttt 1680
cacgtggac agcagggtgtc tcagttcgtc gtaatcgatc aagctggccgg ggttaacacag 1740
tccgtcccta ggggtctt tttccatgtat gtagctccac tcaggcacag acagcgtct 1800
atcgcactca ggatttccca gcagccggcc agcaatagaa caatcgccca gttccagagg 1860
aggaatgcca ttcaagtgtgc acagcttgcg gttgtgggtc ttttccagaa tatccttggc 1920
gtgggtcactg gtcacatttc tttccaggat ggtgtccacc ttctcggtgc tattgtggc 1980
gtggtagccg atacagatct gatcgccccg cacagctgtat aacagcagggta tcaggttagat 2040
gatggccat

```

-continued

<210> SEQ_ID NO 49
 <211> LENGTH: 2064
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 49

atgaaaacca	tcattgcctc	gagctacatc	ttttgtctgg	ctctggccca	ggatctgcc	60
ggcaatgata	atagcaccgc	caccctgtgt	ctgggacacc	acgcgcgtgcc	taatggcacc	120
ctggtgaaaa	ccattaccga	cgaccagatc	gaagtgacca	atgccaccga	gctggtgcag	180
agcagcagca	ccggcaagat	ctgcaacaac	ccccacagaa	tcctggatgg	catcgactgt	240
accctgatcg	atgcctgtct	gggcgatct	cactgcgacg	tgttccagaa	cgagacatgg	300
gacctgttcg	tggagagaag	caaggccttc	agcaactgt	acccctacga	tgtgcggat	360
tacgcctctc	tgagaagcct	ggtggccagc	agcggcacac	tggaaattcat	caccgagggc	420
tttaccttgg	caggcgtgac	ccagaatggc	ggcagcaatg	cctgtaaaag	aggccctggc	480
agcggcttct	ttagcagact	gaactggctg	accaagtccg	gcagcaccta	ccctgtgt	540
aacgtgacca	tgcccaacaa	cgacaacttc	gacaagctgt	acatctgggg	cgtgcaccac	600
cctagcacca	atcaggaaca	gaccagcctg	tacgtgcagg	ccagcggcag	agtgaccgt	660
tctaccagac	ggtcccagca	gaccatcatc	cccaacatcg	agtcaagacc	ttgggtgcgc	720
ggcctgagca	gcagaatcag	catctactgg	accatcgtga	aacctggcga	cgtgtggtg	780
atcaacagca	atggcaaccc	gategcccc	agaggctact	tcaagatgcg	gaccggcaag	840
agcagcatca	tgagaagcga	cgccccatc	gatacctgta	tcagcgagg	catcacc	900
aacggcgcga	tcccaacga	caaggccttc	cagaacgtga	acaagatcac	ctacggcgcc	960
tgccttaagt	acgtgaagca	gaacaccctg	aagctggcca	ccggcatgag	aatgtgccc	1020
gagaaggcaga	caagaggcct	gttggcgcc	attgcccgt	ttatcgagaa	cggtggag	1080
ggcatgatcg	atgggtggta	cggcttcaga	caccagaatt	ctgaggccac	aggacaggcc	1140
gcccgtatcg	agtctacaca	ggccgccc	gaccatcg	acggcaagct	gaacagatgt	1200
atcgagaaaa	ccaacgagaa	gttccaccag	atcgagaaag	aattcagcga	ggtggaggc	1260
agaatccagg	acctggaaaa	atacgtggag	gacaccaaga	tgcacgtgt	gagctacaat	1320
gccgaactgc	tggtcgcct	ggaaaaccag	cacaccatcg	acctgaccga	cagcgagat	1380
aataagctgt	tcgaaaagac	cagacggcag	ctgagagaaa	acgccgagga	catggcaac	1440
ggctgtttca	agatctacca	caagtgcac	aacgcctgca	tgcagagcat	cagaaacggc	1500
acctacgacc	acgtatgtta	cagggacgag	gccctgaaca	acagattcca	gatcaagtcc	1560
ggaggcgcaca	tcatcaagct	gctgaacgag	caggtgaaca	aggagatgca	gagcagcaac	1620
ctgtatcgatcg	gcatgagcag	ctgggtctac	acccacagcc	tggacggcgc	cgccctgttc	1680
ctgttcgacc	acgcccggcga	ggagtacgag	cacgccaaga	agctgtatcat	cttcttgcac	1740
gagaacaacg	tgccctgtca	gctgaccaggc	atcagcgccc	ccgagcacaa	gttcgagg	1800
ctgacccaga	tcttccagaa	ggcctacgag	cacgagcgc	acatcagcga	gagcatcaac	1860
aacatcgtgg	accacgccc	caagagcaag	gaccacgcca	ccttcaactt	cctgcagtgg	1920
tacgtggccg	agcagcacga	ggaggaggt	ctgttcaagg	acatcctgg	caagatcgag	1980
ctgatcgcc	acgagaacca	cgccctgtac	ctggccgacc	agtacgtgaa	gggcacgtcc	2040
aagagcagga	agagcggatc	ctag				2064

US 9,441,019 B2

159**160**

-continued

<210> SEQ ID NO 50
<211> LENGTH: 687
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 50

```

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Phe Cys Leu Ala Leu Gly
1           5          10          15

Gln Asp Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20          25          30

His His Ala Val Pro Asn Gly Thr Leu Val Lys Thr Ile Thr Asp Asp
35          40          45

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
50          55          60

Gly Lys Ile Cys Asn Asn Pro His Arg Ile Leu Asp Gly Ile Asp Cys
65          70          75          80

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro His Cys Asp Val Phe Gln
85          90          95

Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Phe Ser Asn
100         105         110

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
115         120         125

Ala Ser Ser Gly Thr Leu Glu Phe Ile Thr Glu Gly Phe Thr Trp Thr
130         135         140

Gly Val Thr Gln Asn Gly Gly Ser Asn Ala Cys Lys Arg Gly Pro Gly
145         150         155         160

Ser Gly Phe Phe Ser Arg Leu Asn Trp Leu Thr Lys Ser Gly Ser Thr
165         170         175

Tyr Pro Val Leu Asn Val Thr Met Pro Asn Asn Asp Asn Phe Asp Lys
180         185         190

Leu Tyr Ile Trp Gly Val His His Pro Ser Thr Asn Gln Glu Gln Thr
195         200         205

Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr Val Ser Thr Arg Arg
210         215         220

Ser Gln Gln Thr Ile Ile Pro Asn Ile Glu Ser Arg Pro Trp Val Arg
225         230         235         240

Gly Leu Ser Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
245         250         255

Asp Val Leu Val Ile Asn Ser Asn Gly Asn Leu Ile Ala Pro Arg Gly
260         265         270

Tyr Phe Lys Met Arg Thr Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
275         280         285

Pro Ile Asp Thr Cys Ile Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
290         295         300

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Lys Ile Thr Tyr Gly Ala
305         310         315         320

Cys Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
325         330         335

Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Leu Phe Gly Ala Ile Ala
340         345         350

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Ile Asp Gly Trp Tyr Gly
355         360         365

```

-continued

Phe Arg His Gln Asn Ser Glu Gly Thr Gly Gln Ala Ala Asp Leu Lys
370 375 380

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Val
385 390 395 400

Ile Glu Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
405 410 415

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
420 425 430

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
435 440 445

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
450 455 460

Glu Lys Thr Arg Arg Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
465 470 475 480

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Glu Ser
485 490 495

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
500 505 510

Asn Asn Arg Phe Gln Ile Lys Ser Gly Gly Asp Ile Ile Lys Leu Leu
515 520 525

Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser
530 535 540

Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe
545 550 555 560

Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile
565 570 575

Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser
580 585 590

Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala
595 600 605

Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp
610 615 620

His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp
625 630 635 640

Tyr Val Ala Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu
645 650 655

Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala
660 665 670

Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
675 680 685

<210> SEQ ID NO 51
<211> LENGTH: 2064
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 51

ctaggatccg ctcttcctgc tcttggcgat gcccgtcagc tactggtcgg ccaggtacag 60
gccgtggttc tcgttgccga tcagctcgat cttgtccagg atgtccttga acagcacctc 120
ctcctcgatgc tgctcgccca cgtaccactg caggaagtgg aaggtggcgt ggtcccttgct 180
cttgcgtggcg tggccacga tggtgttgc gctctcgctg atgtgtcgct cgtgcgtca 240
ggccttcgg aagatctggg tcaggccctc gaaccttgtc tcggggccgc tgatgtcggt 300

-continued

cagctgcacg ggcacgttgt tctcgtttag gaagatgatc agcttcttgg cgtgctcgta	360
ctccctcgccg gcgtggtcga acaggaaacag gccggcgccg tccaggctgt ggggttagca	420
ccagctgctc atgctcatgt acaggttgct gctctgcata tccctgttca cctgctcgta	480
cagcagcttg atgatgtcgc ctccggactt gatctggaaat ctgttggtaa gggcctcgta	540
cctgtacaca tcgtggtcgt aggtgccgtt tctgatgctc tcatgcagg cggtgtcgca	600
cttggtagtgcg atcttgaagc agccgttgcg catgtcccg cggttttctc tcaatgtcgca	660
tctggctttt tcgaacagct tattcatctc gctgtcggtc aggtcgatgg tttgtgtgtt	720
ttccaggccg accagcagtt cggccattgtt gctccacagg tccatgttgg tttctccac	780
gtattttcc aggttctggta ttctgccttc cacctcgctg aatttttctt cgtatgtgtt	840
gaacttctcg ttggtttctt cgtatctctt gttcagttt ccgttgcattt ggtcgatggc	900
ggccgtgtgtt gacttcagat cggccggccgt tccgtgtcccc tccagaattttt ggtgtctgaa	960
gcgcgtaccac ccattcgatca tgcctccca gccgttctcg ataaaggccgg caatggcgcc	1020
aaacaggccctt cttgtctgtt tctcgccac atttctcatg ccgttggcca gtttccagggt	1080
gttctgttcc acgtacttag ggcaggccgc gtaggtgttcc ttgttccacgt tctggaaagg	1140
cttgcgtttt gggatgtgtc cggtgggggtt gatgcactcg ctgatacagg tatcgatgg	1200
ggcgctegctt ctcatgtgtc tgctcttgcg ggtccgcattt ttgaagttagc ctctggggcc	1260
gatcagggttgc cattgtgtt tgatcaccag cacgtcgccca gggttccacga tggccacgtt	1320
gatgctgttgc ctgctgtca ggccgcgcac ccaaggctt gactcgatgt tggggatgtat	1380
ggtctgttgg gaccgtctgg tagacacggcgtt cactctgcgtt ctggccgtca cgtacaggct	1440
ggtctgttcc tgattgggtc taggggtggcgtt cacggccccc atgtacatgt tggcgtttttt	1500
gtcggttgc ggcattgtca cggttgcac agggtaggtt ctggccggact tggcgttgc	1560
gttcagtctg ctgaagaagc cgctgcagg gccttttta caggcatttc tgccgcattt	1620
ctgggttcacg cctgtccagg taaaaggccctt ggtgtatgtt tccagtgtgc cgctgtggc	1680
caccaggctt ctcaagaggcgtt cgtatcggtt cacatcgtag gggtagcgtt tgctgaaggc	1740
cttgcgttctc tccacgaaca ggtcccatgtt ctgcgttgcgtt aacacgtcgatgtt gaggatgt	1800
gcgcgcaggcgc gcatcgatca gggtagcgtt gatgcacatcc aggattctgtt ggggggtttttt	1860
gcagatcttgc cccgtgtgtc tgctctgcac cagctcggtt gcattgggtca ctgcgtatgt	1920
gtcggttgcgtt atggtttca ccagggtgtcc attaggcactt ggtgtgttgc ccagacacag	1980
ggtggccgggtt ctattatcat tgccggccag atccctggccc agagccagac aaaagatgtt	2040
gctcaggccatgtt tcat	2064

<210> SEQ_ID NO 52
 <211> LENGTH: 2064
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 52

atgaaaaacca tcattggccct gagctacatc ctgtgcctgg tggccacaca gaagctggcc	60
ggcaacgata atagcaccgc cacactgtgtt ctggcacacc acggccgtgcc taatggcacc	120
atcgtaaaaaa caatcaccaa cgaccagatc gaagtgcacca atgcccacaga gctgggtcag	180
agcagcagca cagggcagatc ctgtgacagc cccaccaga tccctggatgg cgagaactgt	240

-continued

accctgatcg atgcctgct ggccatct cagtgcacg gttccagaa caagaatgg	300
gacctgtcg tggagagaag caaggctac agcaactgct acccctacga cgtgcctgat	360
tacgccagcc tgagaaggct ggtggctct agcggcaccc tggattcaa caacgagac	420
ttcaactgga ccggctgac acagaatggc accagcagcg cctgcacag acggtccaac	480
aacagttct ttagtagact gaattggctg acccactga agttcaagta cccgcctg	540
aacgtgacca tgcccaacaa tgagaagttc gacaagctgt acatctgggg agtgcaccac	600
cctggcaccc acaacgatca gatcttcct tacgcccagg ccagcggcag aatcaccgtg	660
tccccaaga gaagccagca gaccgtgatc cccatatcg gcagcagacc cagagtgcgg	720
aacatccccca gcaggatcg catctactgg acaatcgta agcctggcga catcctgctg	780
atcaacgca ccggcaacct gatgcctct cggggctact ttaagatcg aagcggcaag	840
agcagcatca tgagatccga cgccccatc ggcaagtgc acagcggatg catcaccacca	900
aacggcagca tcccaacga caagcccttc cagaacgtga acaggatcac ctacggcgcc	960
tgcctcatat acgtgaagca gaacaccctg aagctggcca cccgcattgaa aatgtgccc	1020
gagaaggcaga ccagaggcat ctttggcgcc attgcccgt ttatcgagaa tggctggag	1080
ggaatggtgg atgggtggta cggcttcaga caccagaata gcgaggaaat tggacaggcc	1140
gcccgtatcg aatctaccca ggccgcctc gaccagatca acggcaagct gaacaggctg	1200
atcggcaaga ccaacgagaa gttccaccag atcgagaaat aattcagcga ggtggaggc	1260
agaatccagg acctggaaaa atacgtggag gacaccaaga tcgacctgtg gagctacaat	1320
gccgaactgc tggcgcctt ggaaaaccag cacacaattt atctgacaga cagttagat	1380
aataagctgt tcgagaaaac caagaagcag ctgagagaaa acgcccggaa catggcaac	1440
ggctgttca agatctacca caagtgcac aacgcctgc tcggcgcatt cagaaacggc	1500
acctacgacc acgacgtgtc cagagatgag gccctgaaca accggtttca gatcaagtcc	1560
ggaggcgcaca tcatcaagct gctgaacgcg caggtgaaca aggatgcg gagcgcac	1620
ctgtacatga gcatgagcag ctggcttac acccacagcc tggacggcgc cggcctttc	1680
ctgttcgacc acggcccgaa ggagtacgcg cacgcaaga agctgtatcat cttctgaac	1740
gagaacaacg tggcgttca gctgaccgc atcagcgcctt ccgagcacaa gttcgaggc	1800
ctgacccttca tttccatggaa ggcctacgcg cacgagcgc acatcagcga gagcatcaac	1860
aacatcgatgg accacgcctt caagagcag gaccacgcgc cttcaactt cttgcgttgg	1920
tacgtggccg agcagcacga ggaggagggtg ctgttcaagg acatctggaa caagatcgag	1980
ctgatcggca acgagaacca cggcctgtac ctggccgcacc agtacgtgaa gggcatcgcc	2040
aagagcggaa agagcggatc ctat	2064

<210> SEQ ID NO 53
 <211> LENGTH: 687
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 53

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Thr			
1	5	10	15

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly			
20	25	30	

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp

US 9,441,019 B2

167

-continued

168

35	40	45
Gln Ile Glu Val Thr Asn Ala Thr	Glu Leu Val Gln Ser Ser Ser Thr	
50	55	60
Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Glu Asn Cys		
65	70	75
Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln		
85	90	95
Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn		
100	105	110
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val		
115	120	125
Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr		
130	135	140
Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Asn		
145	150	155
Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Lys Phe Lys		
165	170	175
Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys		
180	185	190
Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Asn Asp Gln Ile		
195	200	205
Phe Pro Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg		
210	215	220
Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Arg Val Arg		
225	230	235
Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly		
245	250	255
Asp Ile Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly		
260	265	270
Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala		
275	280	285
Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile		
290	295	300
Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala		
305	310	315
Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met		
325	330	335
Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala		
340	345	350
Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly		
355	360	365
Phe Arg His Gln Asn Ser Glu Gly Ile Gly Gln Ala Ala Asp Leu Lys		
370	375	380
Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu		
385	390	395
Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser		
405	410	415
Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr		
420	425	430
Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu		
435	440	445
Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe		
450	455	460

US 9,441,019 B2

169**170**

-continued

Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
 465 470 475 480
 Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
 485 490 495
 Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
 500 505 510
 Asn Asn Arg Phe Gln Ile Lys Ser Gly Gly Asp Ile Ile Lys Leu Leu
 515 520 525
 Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser
 530 535 540
 Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe
 545 550 555 560
 Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile
 565 570 575
 Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser
 580 585 590
 Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala
 595 600 605
 Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp
 610 615 620
 His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp
 625 630 635 640
 Tyr Val Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile Leu
 645 650 655
 Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala
 660 665 670
 Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 675 680 685

<210> SEQ ID NO 54
 <211> LENGTH: 2064
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

 <400> SEQUENCE: 54

```

ctaggatccg ctttcctgc tcttggcgat gcccggcact tactggtcgg ccaggtacag 60
gccgtggttc tcgttgccga tcagctcgat cttgtccagg atgtccttga acagcacctc 120
ctcctcgtgc tgctcgccca cgtaccactg caggaagttg aagggtggcgt ggtccttgc 180
cttgcgtggcg tggtccacga tggatgttgcat gctctcgctg atgtgcgtc cgtgcgtca 240
ggccttctgg aagatctggg tcaggccctc gaacttgtgc tcggggccgc tgatgtgg 300
cagctgcacg ggcacgttgt tctcggtcag gaagatgatc agcttcttgg cgtgcgtca 360
ctcctcgccg gcgtggtcga acaggaacag gccggccgcg tccaggctgt ggggttagca 420
ccagctgctc atgctcatgt acaggttgct gctctgcata tccctgttca cctgcgtc 480
cagcagtttgcg atgatgtcgc ctccggactt gatctgaaac cggttgcgtca gggccctc 540
tctgtacacg tcgtggtcgt aggtggccgtt tctgtatgcgtc cccatgcagg cgttgtcgca 600
cttgggtttc tctggatgttgc catgttgcgtc ggttttgcgtca tctgtatgcgtc 660
ttccaggccg accagcagtt cggcattgtta gctccacagg tcgtatcttgg tgcgttgc 720
ttccaggccg accagcagtt cggcattgtta gctccacagg tcgtatcttgg tgcgttgc 780
  
```

-continued

gtatTTTcC aggtccTgga ttctgcCCTC caccTcgTg aattcttCTC cgatctggTg 840
gaacttctcg ttggTcttgc cgatcagCCT gttcagTgtg cegttgatct ggTcgatggC 900
ggcTgggta gatttcagat cgggggcTg tccaaTTccc Tcgctattct ggtgtctgaa 960
gccgtaccaC ccATCCACCA ttccCTCCCA gccattctcg ataaAGCCGG caatggcgCC 1020
aaagatgcct ctggTctgct tctcgggcac atttctcatg ccggTggCCA gcttcagggt 1080
gttctgttC acgttatctag ggcaggcgCC gtaggtgatc ctgttCACGT tctggaaaggG 1140
cttgcgttg gggatgctgc cgTTggggT gatgcactcg ctgttgcact tgccgatggg 1200
ggcgtcgat ctcATgtatgc tgcttgcAC tcttctgatc taaaagttagc cccgaggggc 1260
gatcagggttC ccggTgtctgt tgatcagcag gatgtcgcca ggTTcacga ttgtccagta 1320
gatgctgatc ctgtggggA tgTTccgcaC tctgggtctg ctgcccataT tgggatcac 1380
ggTctgttg ctTctttgg tggacacggT gattctgcgg ctggcctggg cgtaaaggaa 1440
gatctgatcg ttgtcggtgc cagggtggTg cactccccag atgtacagat tgcgaaACTT 1500
ctcattgttg ggcATggtca cgttcaGGGC ggggtacttg aacttcaggT gggTcagCCA 1560
attcagtctA ctgaagaAGC tgTTgttggA ccgtctgatc caggcgtgc tggtgcCATT 1620
ctgtgtcAcC ccggTccAGT tgaagctctc gttgttgaat tccagggtgc cgctagaggc 1680
caccaggctt ctcaggctgg cgtaatcagg cacgtcgtag gggtagcagt tgctgttaggc 1740
cttgcttctc tccacgaaca ggtcccattt ctgttctgg aagccgtcgC actgaggatc 1800
geccagcagg gcATcgatca gggTacagtt ctgcCcatcc aggatctggT gggggctgtc 1860
acagatctcg cctgtgtcgc tgctctgcac cagctctgtg cattggTca ctgcAtctg 1920
gtcgTTgggtg attgtttca cgtatggTgcC attaggacacg cgcgtggTgc ccagacacag 1980
tgtggcggtg ctattatcgt tgccgggcag ctTctgttg aacaccaggc acaggatgtA 2040
qctcaaggqca atqatggtt tcat 2064

```
<210> SEQ ID NO 55
<211> LENGTH: 2067
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 55

atggaaaaga tcgtgctgct	gctggccatt gtgagcctgg tgaagagcga ccagatctgc	60
atgggctacc acgccaacaa tagcacagag cagggtggaca ccatcatggaa	aaaaaacgtg	120
accgtgaccc acgctcagga catcctggaa aagaccacaa acggcaagct gtgtgatctg		180
gacggcgctga agcctctgat cctgagagat tgtagcgtgg ctggatggct gctgggcaac		240
cctatgtgcg acgagttcat caacgtgccc gagtgagact atatcgtgga gaaggccaac		300
ccccccaacg atctgtgtta ccccgccgc ttcaacgatt acgaggaaact gaagcacctg		360
ctgtcccgga tcaaccacctt cgagaagata cagatcatcc ccaagtcttc ttggagcgat		420
cacgaagccct ctagcgaggat gtctagcgcc tggccttacc tggcagcccc cagttcttc		480
agaaaacgtgg tgtggctgat caagaagaac agcacctacc ccaccatcaa gaagagctac		540
aacaacacca accaggaaga tctgctggtc ctgtggggaa tccaccaccc taatgatgcc		600
gcccggcggc ccagactgta ccagaaccccc accacctata ttagcatgg caccggcacc		660
ctgaatcaga gactggtgcc caagatcgcc accagatcca aggtgaacgg ccagagcgcc		720
aggatggaat tcttctggac catcctgaag cccaaacggacg ccatcaactt cgagagcaac		780

-continued

```

ggcaacttta tcgcccctga gtacgcctac aagatcgtga agaaggcgaa cagcgccatc      840
atgaagagcg agctggaaa cggcaactgc aacaccaagt gccagacacc tatggcgcc      900
atcaacagca gcatgccctt ccacaacatc caccctctga ccatacgccgaa gtgccttaag      960
tacgtgaaga gcaacagact ggtgtggcc acaggcctga gaaatagccc ccagcgggag      1020
agcagaagaa agaagagggg cctgtttgga gccatgcggc gctttattga aggccggctgg      1080
cagggatgg tggatggctg gtacggctac caccacagca atgagcaggc ctctggatata      1140
gccgcgacaa aagagtctac ccagaaggcc atcgacggcg tcaccaacaa ggtgaacagc      1200
atcatcgaca agatgaacac ccagttcgag gctgtggccgagatgtcaa caacctggaa      1260
cggcggatcg agaacctgaa caagaaaatg gaagatggctt ttctggatgt gtggacctac      1320
aatgcccgaac tgctgggtct gatggaaaac gagcggacc tggacttcca cgacagcaac      1380
gtgaagaacc tgtaacgacaa agtgccggctg cagctgagag acaacgccaa agagctggc      1440
aacggctgtc tgcgttctt ccacaagtgc gacaacgagtc gatggaaag catcggaaac      1500
ggcacctaca actaccctca gtacagcgag gaagccaggc tgaagaggga agagatcagc      1560
tccggaggcg acatcatcaa gctgtgaac gagcagggtg acaaggagat gcagacgc      1620
aacctgtaca tgagcatgag cagctggtc tacacccaca gctggacgg cgccggctg      1680
ttcctgttcg accaegccgc cgaggagttac gggcacccca agaagctgtat catcttctg      1740
aacgagaaca acgtgcccgt gcagctgacc agcatcagcg ccccgagca caagttcgag      1800
ggcctgaccc agatcttca gaaggcctac gggcacccgagc agcacaatcg cgagacatc      1860
aacaacatcg tggaccacgc catcaagagc aaggaccacg ccaccttcaa cttcctgcag      1920
tggtaacgtgg ccgagcagca cgaggaggag gtgctgttca aggacatct ggacaagatc      1980
gagctgatcg gcaacgagaa ccacggcctg tacctggccg accagtagt gaagggcatc      2040
gccaagagca ggaagagcgatcctag                                         2067

```

<210> SEQ ID NO 56

<211> LENGTH: 688

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 56

Met	Glu	Ile	Val	Leu	Leu	Ala	Ile	Val	Ser	Leu	Val	Lys	Ser
1				5			10			15			

Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val
				20			25			30					

Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile
			35			40			45						

Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys
				50			55		60						

Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn
65				70			75		80						

Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val
			85			90		95							

Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Ser	Phe	Asn
			100			105		110							

Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu
				115			120		125						

-continued

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Asn Ser Thr Tyr Pro Thr Ile
165 170 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
195 200 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
210 215 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
260 265 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305 310 315 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
325 330 335

Pro Gln Arg Glu Ser Arg Arg Lys Arg Gly Leu Phe Gly Ala Ile
340 345 350

Ala Gly Phe Ile Glu Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
355 360 365

Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
370 375 380

Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
385 390 395 400

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
405 410 415

Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
420 425 430

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
435 440 445

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
450 455 460

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465 470 475 480

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
485 490 495

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510

Arg Leu Lys Arg Glu Glu Ile Ser Ser Gly Gly Asp Ile Ile Lys Leu
515 520 525

Leu Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met
530 535 540

Ser Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu

-continued

545	550	555	560
Phe Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu			
565	570	575	
Ile Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile			
580	585	590	
Ser Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys			
595	600	605	
Ala Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val			
610	615	620	
Asp His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln			
625	630	635	640
Trp Tyr Val Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile			
645	650	655	
Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu			
660	665	670	
Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser			
675	680	685	

<210> SEQ_ID NO 57
<211> LENGTH: 2067
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 57

ctaggatccg ctttcctgc tcttggcgat gccc ttca cg tactggtcgg ccaggta cag	60
gccgtgggtc tcgttgcgc ta ca gctcgat cttgtccagg atgtccttga ac agc ac ct	120
c t c c t c g t g c tg ct cggccca cgt acc act g c agg a a g t t g c g t g g t c	180
ctt gatggcg tgg tccacga t g t g t g t g a t g t c g t g t at g t g t g t g a	240
ggc tt ct gg a a g a t ct gg gg t c a g g c c c t c g a a c t t g t g c t c g g g	300
c a g c t g c a c g g g c a c g t t g t c t c g t t c a g a g t g a t c a g t t c t t g	360
c t c c t c g g c g g c t g g t c g a a c g g a a c a g g c g c c g t c a g g c t g t g	420
c c a g c t g c t c a t g c t c a t g t a c a g g t t g t g t c t g c a t c t g c t c	480
c a g c a g t t g t a c g a t g t c g c t c c c g g a g a c t g t t c t c t t c	540
g c t g t a c t g a g g t a g t t g t g t g a g g t g c c g t t c t g a t g c a c t	600
c t t g t g g t a g a c t c g a a g c a g c t t g t g c a g t t g t c t c t c	660
c c g c a c t t g t c g t a c a g g t t g t t c t c t c t c t c t c t c t c t c	720
t t c c a t c a g c a c c a g a c t t g t a g t c c a c a c a t c a g g a a g c	780
t t t c t t g t c a g g t t c t c a g g t t g t t g a a c t c t c t g c a c a g c	840
g a a c t g g g t g t c t c a t c t c t c t c t c t c t c t c t c t c t c t c t c	900
c t t c t g g g t a g a c t t t g t c g g g g c a t c a g a g c c c t c t c t c	960
g c c g t a c c a g c a c c a t t c c t c t c t c t c t c t c t c t c t c t c t c	1020
a a a c a g g g c c c t c t t c t c t c t c t c t c t c t c t c t c t c t c t c t c	1080
c a g c a c c a g t t c a c t c t c t c t c t c t c t c t c t c t c t c t c t c t c	1140
g t t g t g g a a g g c a t g c t c g t g t g a t g c a c t g t g a c t g t g t g	1200
g t t g c c g t a t c c a g c t c g c t c t c t c t c t c t c t c t c t c t c t c	1260
g g c g t a c t c a g g g c g a t a a g t g c c g t t c a g t t c g a a g t g g c	1320

-continued

caggatggtc cagaagaatt ccatacctgcc gctctggccg ttcaccttgg atctgggtgc	1380
gatcttgggc accagtctct gattcagggt gctggtgcgg atgctgatata aggtgggtgg	1440
gttctggta cagtctggtct gctcgccggc atcattaggg tggtgattc cccacaggac	1500
cagcagatct tcctgggtgg tggtgttga gcttttttg atgggggggtt aggtgttgtt	1560
cttcttgatc agccacacca cgtttctgaa gaagctgggg ctgcccaggta aaggacaggc	1620
gctagacact ccgcttagagg ctctcgatc gctccaagag gacttggggta tgatctggat	1680
cttctcgaag tggttgatecc gggacacgag gtgtttcaat ttctcgtaat cggttgaagct	1740
gccccggtaa cacagatctg tgggggggtt ggcccttc acgatatacg tccactcggg	1800
cacgttcatg aactcgctgc acatagggtt gcccagcaggc catccagccca cgctacaatc	1860
tctcaggatc agagggttca cgccgtccag atcacacaggc ttggcggtgtt ggggttttc	1920
caggatgtcc tgagegtggg tcacggtcac gttttttcc atgatgggtt ccacctgctc	1980
tgtgttatgg ttggcggtgtt agccaatgtca gatctggtcg ctttcacca ggctcacaat	2040
ggccagcaggc agcacatct ttccat	2067

<210> SEQ ID NO 58

<211> LENGTH: 2109

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 58

atgaaggcca tcategtgct gctgtatggtg gtgaccagca acgcccatacg aatctgcacc	60
ggcatcacca gcagcaatacg ccccccattgtgt gtaaaaacacg ccacccaggcg cgaagtgaat	120
gtgacaggcg tgatccctct gaccaccacc cccaccaaga gctacttcgc caacctgaag	180
ggcaccaggaa ccagaggcaa gctgtggccc gattgcctgaa actgcaccga tctggatgt	240
gtctctggca gacatatgtt tggggcacc acaccatctg ccaaggccag catctgcac	300
gaagtgaagc ctgtgaccag cggtgtcttc cccatcatgc acgacccggac caagatcaga	360
cagctgcccac acctgtctgaa aggttacgag aacatccggc tgtccaccca gaatgtgtat	420
gtatgcggaga aagccctgg cggacattat agactggca ccagcggctc ttgtcccaat	480
gccacctcca agageggctt tttgccaca atggcctggg ccgtgectaa ggacaacaac	540
aagaacgcca ccaaccctct gaccgtggag gtgccttaca tctgtacaga gggcgaggat	600
cagatcacag tgtggggctt ccacagcgac gacaagaccc agatgaagaa cctgtacggc	660
gacagcaacc cccagaagtt taccagcaggc gccaatggcg tgaccaccca ctacgtgtcc	720
cagatcgccca gctttcccgat tcagacagag gatggcggac tgcctcagtc tggcaggatc	780
gtgggtggact acatgtatgc gaaacgtggc aagacccggc ccatcgatc tcagagaggc	840
gtgtgtgtgc ctcagaaagt gtgggtgtcc agcggcaggat ctaaaatgtat caagggcaggc	900
ctgcctctga ttggcgaggc cgactgtctg cacgaaaatg acggcggcct gaacaagagc	960
aagccctact acacaggcgaa gcacgccaag gccatcgccaa attggcccat ctgggtgaaa	1020
accccccgttga agctggccaa tggcaccaag tacagacctc ccgccaagct gctgaaagag	1080
agaggcttct ttggcgccat tgccggattt ctggaaaggcg gctggggaggaa atgattgcc	1140
ggctggcaggc gctatacatac tcatggggcc catggcgtgg ctgtggccgc cgatctgaag	1200
tctaccagg aagccatcaa caagatcacc aagaacctga acagcctgag cgagctggaa	1260

-continued

gtgaagaatc	tgcagagact	gagggcgccc	atggatgagc	tgcacaacga	gatcctggaa	1320
ctggacgaga	aagtggatga	tctccgcgcc	gatacaattt	cctcccatat	tgaactggcc	1380
gtgctgtgt	ccaaacgaggg	catcatcaac	agcgaggatg	aacacctgtc	ggccctggaa	1440
cggaagctga	agaagatgct	gggccttct	gccgtggaga	tccgcaacgg	ctgcttcgag	1500
acaaaggaca	agtgcaccca	gaccgtgcctg	gatagaatcg	ccgctggcac	cttcaatgcc	1560
ggcgagttca	gcctgcctac	cttcgacagc	ctgaatatca	cctccggagg	cgacatcatc	1620
aagctgctga	acgagcagg	gaacaaggag	atgcagagca	gcaacctgt	catgagcatg	1680
agcagctgg	gctacaccca	cagcctggac	ggcgccggcc	tgttccctgtt	cgaccacgcc	1740
gccgaggagt	acgagcacgc	caagaagctg	atcatcttcc	tgaacgagaa	caacgtgccc	1800
gtcagctga	ccagcatcag	cggccccgg	cacaagttc	agggcctgac	ccagatctc	1860
cagaaggcct	acgagcacga	gcagcacatc	agcgagagca	tcaacaacat	cgtggaccac	1920
gccccatcaaga	gcaaggacca	cgccacatc	aacttcctgc	agtggtagt	ggccgagcag	1980
cacgaggagg	aggtgctgtt	caaggacatc	ctggacaaga	tccgagctgt	cgccaacgag	2040
aaccacggcc	tgtacacccgc	cgaccagtac	gtgaaggcga	tccgccaagag	caggaagagc	2100
ggatccttag						2109

<210> SEQ_ID NO 59

<211> LENGTH: 702

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 59

Met	Lys	Ala	Ile	Ile	Val	Leu	Leu	Met	Val	Val	Thr	Ser	Asn	Ala	Asp
1					5			10			15				

Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys
					20			25			30				

Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Thr
						35		40			45				

Thr	Thr	Pro	Thr	Lys	Ser	Tyr	Phe	Ala	Asn	Leu	Lys	Gly	Thr	Arg	Thr
						50		55			60				

Arg	Gly	Lys	Leu	Cys	Pro	Asp	Cys	Leu	Asn	Cys	Thr	Asp	Leu	Asp	Val
						65		70			75			80	

Ala	Leu	Gly	Arg	Pro	Met	Cys	Val	Gly	Thr	Thr	Pro	Ser	Ala	Lys	Ala
						85		90			95				

Ser	Ile	Leu	His	Glu	Val	Lys	Pro	Val	Thr	Ser	Gly	Cys	Phe	Pro	Ile
						100		105			110				

Met	His	Asp	Arg	Thr	Lys	Ile	Arg	Gln	Leu	Pro	Asn	Leu	Arg	Gly	
						115		120			125				

Tyr	Glu	Asn	Ile	Arg	Leu	Ser	Thr	Gln	Asn	Val	Ile	Asp	Ala	Glu	Lys
						130		135			140				

Ala	Pro	Gly	Gly	Pro	Tyr	Arg	Leu	Gly	Thr	Ser	Gly	Ser	Cys	Pro	Asn
						145		150			155			160	

Ala	Thr	Ser	Lys	Ser	Gly	Phe	Phe	Ala	Thr	Met	Ala	Trp	Ala	Val	Pro
						165		170			175				

Lys	Asp	Asn	Asn	Lys	Asn	Ala	Thr	Asn	Pro	Leu	Thr	Val	Glu	Val	Pro
						180		185			190				

Tyr	Ile	Cys	Thr	Glu	Gly	Glu	Asp	Gln	Ile	Thr	Val	Trp	Gly	Phe	His
						195		200			205				

-continued

Ser Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn Pro
 210 215 220
 Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
 225 230 235 240
 Gln Ile Gly Ser Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
 245 250 255
 Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
 260 265 270
 Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
 275 280 285
 Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
 290 295 300
 Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
 305 310 315 320
 Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
 325 330 335
 Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
 340 345 350
 Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala
 355 360 365
 Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly
 370 375 380
 Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys
 385 390 395 400
 Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu
 405 410 415
 Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp
 420 425 430
 Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu
 435 440 445
 Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser
 450 455 460
 Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu
 465 470 475 480
 Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly Asn
 485 490 495
 Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg
 500 505 510
 Ile Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe
 515 520 525
 Asp Ser Leu Asn Ile Thr Ser Gly Gly Asp Ile Ile Lys Leu Leu Asn
 530 535 540
 Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met
 545 550 555 560
 Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu
 565 570 575
 Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile
 580 585 590
 Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala
 595 600 605
 Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr
 610 615 620
 Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His

-continued

625	630	635	640
Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr			
645	650	655	
Val Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile Leu Asp			
660	665	670	
Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp			
675	680	685	
Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser			
690	695	700	

<210> SEQ_ID NO 60
<211> LENGTH: 2109
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 60

ctaggatccg ctcttcctgc tcttggcgat gccc ttca cg tactgg tcgg ccaggta cag	60
gccgtgggtc tcgttgcga tc agtcgtcat cttgtccagg atgtccttga acagcacctc	120
ctccctcgtgc tgctcggcca cgtaccactg caggaagt tg aagg tggc gt ggtc ttgct	180
cttgatggcg tgg tccacga tgg tttgat gcttcgctg atgtgctgct cgtgctcgta	240
ggccttctgg aagatctggg tcaggccctc gaacttgc tcggggccgc tgatgtgg	300
cagctgcacg ggcacgttgc tctcgttca gaa agatgatc agcttcttg cgtgctcgta	360
ctccctcggcg gcgtggtcga acaggaacag gcccggccgc tccaggctgt gggtagca	420
ccagctgctc atgctcatgt acagg tttgc gcttcgatc tcc ttgttca cctgctcgta	480
cagcagcttgc atgatgtcgc ctccggaggt gatattcagg ctgtcgaagg taggcaggct	540
gaactcgcgc gcattgaagg tgccagcggc gattctatcc agg caggctgt ggtgcactt	600
gtgctttgtc tcgaaggcgc cgttgcgtat ctccacggca gaagg gcca gcatcttctt	660
cagcttcgttccagg gcaagg ttttc gatc tccatcg ttcgtatgc cctcgttga	720
cagcagcactt cccaggatct ctttgcgtat ctcatccatg gcccgcgtca gtctctgcag	780
attcttcaact tccagctcgc tcaggctgtt cagg ttttgcgtat gtatgttgc ttttgcgtt	840
ctggtagac ttcagatcgg cggccacagg cacgcctatgg gccccatgag atgtatagcc	900
gtgccagccg gcaatcattt cctcccgatc gccttccaga aatccggcaa tggcgccaaa	960
gaaggccttc tcttcagca gcttggcgaa aggtctgtac ttgggtccat tggccagctt	1020
cagggggtt ttcaccaga tggggcaatt gccgatggc ttggcggtc cgcctgtgt	1080
gtagggttttgc tcttgcgtat ggcggccgtat ctttgcgtat agacagtcgg cctcgttca	1140
cagaggcagg ctgccttgc tcactttaga cctgcgtatc gcacaccaca ctttgcgtt	1200
cagcagcactt cctctctgtat acacgatggt gcccgttttgc caggcttctt gcatcatgt	1260
gtccaccacac atcctgcacatc actgaggcgc tccgcctatcc tctgtctgtat cgggaaagct	1320
gccgatctgg gacacgttagt ggggtggcgc gccatttgcgtat ctttgcgttac accttgcgtt	1380
gttgcgtatgc cctgtacaggat tcttcatctg ggtcttgcgtat ctttgcgttac agccccacac	1440
tgtgatgtatgc tcttcgcctt ctgtacagat gttagggcacc tccacggcata gaggggttgcgtt	1500
ggcgttcttg ttgttgttccat taggcacggc ccaggccattt gttggcaaaaa agccgctt	1560
ggaggtggca ttgggacaag agccgcttgcgtat ctttgcgttac taagg tccgc caggggctt	1620
ggaggtggca ttgggacaag agccgcttgcgtat ctttgcgttac taagg tccgc caggggctt	1680

ctcggcatcg atcacattct gggtggacag ccggatgttc tcgttagccctc tcagcaggtt	1740
gggcagctgt ctgatcttgg tccggtcgtg catgatgggg aagcagccgc tggtcacagg	1800
cttcacttcg tgcaggatgc tggccttggc agatggtgtg gtgcccacac acataggct	1860
gcccagagcc acatccagat cggtgcaagtt caggcaatcg gggcacagct tgcctctgg	1920
tctggtgcctt ttcaggttgg cgaagtagct cttgggtggg gtgggtgtca gagggatcac	1980
gctgtcaca ttcaactcgc cctgggtggc tggtttcacc acatgggggc tattgctgct	2040
ggtgatgccc gtgcagattc tatggcggtt gctggteacc accatacga gcacgatgat	2100
ggccttcat	2109

<210> SEQ ID NO 61
<211> LENGTH: 2064
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 61	
atgaaaacca taattgcgtc gtcctacata ctgtgtctgg tgtttgccca gaaactgccc	60
ggcaatgaca actcaacagc cacgctctgc ttggggcacc atgccgtccc taacgggacc	120
attgtaaaaa ccattactaa cgatcagata gaggtgacta atgccaccga gctggtgcaa	180
agtagctcca caggagagat ctgcgatagt ccccaccaga ttctggacgg aaagaattgt	240
acgctgatcg acgcgcgtgtt gggegaccct cagtgtgacg gatttcagaa taagaagtgg	300
gatctgtttt tggaaaggta aaaggcttat tcaaattgct acccttacga tgtgcctgat	360
tatgccagcc tgcggccct cgtcgctct agtgggactc tggagttcaa caacgagtca	420
tttaactgga ctggcggttac acagaacggg actagttccg cttgcataag gagaagcaaa	480
aatagttct tcagecagact gaattggctg acacatctga acttcaagta ccctgcactg	540
aatgtaaacca tgcccaacaa cgagcagttc gataagctt acatttggg agttcatcat	600
cctggcactg acaaggatca gatcttctg tatgcccagg cttccggcag gattaccgtg	660
tctacaaaga gaagccagca aactgtgtct cccaatatcg gcagtagacc cagagtacgg	720
aacatecccta gtcgcatacg tatttactgg accatctgta aaccaggcga tattctctg	780
attaacagta ctggcaacct gategcffff cggggatact ttaaaatccg ctctggaaag	840
tctccattt ttagatcaga tgcaccgtc ggaaaatgca actctgagtg tatcacaccc	900
aatgggagca ttcccaatga caaaccttc cagaacgtta atcgaataac ttatgggcc	960
tgtccacggt acgtgaagca aaataccttg aaactggcga ccggtatgctg caatgtcccc	1020
gaaaaacaga cccgcggat atttgggct atcgcaggct ttatcgagaa tggctggaa	1080
ggatgggtgg atgggtggta tggttttaga catcaaaact ccgaaggcag aggccaggct	1140
gccgatctca agagcacgca ggccgctata gatcagatca atggaaagct caacagactg	1200
atcggaaaaa ccaacgaaaa attccatcg atcgagaaag agttctccga agtcgaggg	1260
cgcatacagg acctggagaa gtatgttgcgata cacaatggcga ttgcgtatgtg tgcctacaat	1320
gccgagctgc tggggctct ggagaatcg cacactattg acctgaccga ttcagagatg	1380
aacaaacttt ttgagaagac gaagaagcag ctttagagaaa atgcagagga catggggAAC	1440
ggatgcttta aaatatatca taagtgtat aatgcctgca tcggatcaat tagaaatgg	1500
acctatgatc acgtatgttta cagggacgaa ggcgtgaata acaggttcca gataaaatcc	1560

US 9,441,019 B2

189**190**

-continued

```

ggaggcgaca tcatcaagct gctgaacgag caggtgaaca aggagatgca gagcagcaac 1620
ctgtacatga gcatgaggcag ctggtgctac acccacagcc tggacggcgc cggccttttc 1680
ctgttcgacc acgcccggca ggagtacgag cacgccaaga agctgatcat cttcctgaac 1740
gagaacaacg tgcccttgca gctgaccgc atcagcgccc ccgagcacaa gttcgagggc 1800
ctgaccaga tcttccagaa ggcctacgag cacgaggcag acatcagcga gagcatcaac 1860
aacatcgtagg accacgccc caagagcaag gaccacgcca cttcaactt cctgcagtgg 1920
tacgtggccg agcagcacga ggaggagggtg ctgttcaagg acatcctgga caagatcgag 1980
ctgatcggca acgagaacca cggcctgtac ctggccgacc agtacgtgaa gggcatcgcc 2040
aagagcagga agagcggatc ctag 2064

```

```

<210> SEQ_ID NO 62
<211> LENGTH: 687
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```
<400> SEQUENCE: 62
```

```

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Ala
1 5 10 15

```

```

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20 25 30

```

```

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
35 40 45

```

```

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
50 55 60

```

```

Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Lys Asn Cys
65 70 75 80

```

```

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
85 90 95

```

```

Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
100 105 110

```

```

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
115 120 125

```

```

Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
130 135 140

```

```

Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Lys
145 150 155 160

```

```

Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Asn Phe Lys
165 170 175

```

```

Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Gln Phe Asp Lys
180 185 190

```

```

Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Lys Asp Gln Ile
195 200 205

```

```

Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
210 215 220

```

```

Ser Gln Gln Thr Val Ser Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
225 230 235 240

```

```

Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
245 250 255

```

```

Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
260 265 270

```

-continued

Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
275 280 285

Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
290 295 300

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
305 310 315 320

Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
325 330 335

Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala
340 345 350

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
355 360 365

Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala Ala Asp Leu Lys
370 375 380

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu
385 390 395 400

Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
405 410 415

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
420 425 430

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
435 440 445

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
450 455 460

Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
465 470 475 480

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
485 490 495

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
500 505 510

Asn Asn Arg Phe Gln Ile Lys Ser Gly Gly Asp Ile Ile Lys Leu Leu
515 520 525

Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser
530 535 540

Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe
545 550 555 560

Leu Phe Asp His Ala Ala Glu Glu Tyr His Ala Lys Lys Leu Ile
565 570 575

Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser
580 585 590

Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala
595 600 605

Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp
610 615 620

His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp
625 630 635 640

Tyr Val Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile Leu
645 650 655

Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala
660 665 670

Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
675 680 685

-continued

<210> SEQ_ID NO 63
 <211> LENGTH: 2064
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 63

ctaggatccg	ctttcctgc	tcttggcgat	gcccttcacg	tactggtcgg	ccaggtacag	60
gcccgtggtc	tcgttgcga	tcagctcgat	cttgcagg	atgtccgt	acagcacctc	120
ctcctcggtc	tgctcgcca	cgtaccactg	caggaagtgg	aagggtggcgt	ggtccttgc	180
cttgcgtggc	ttgtccacga	tgttgttgc	gtctcgctg	atgtgtgc	cgtgtcgta	240
ggccttcgttgc	aagatctggg	tcaggccctc	gaacttgc	tgcggggcgc	tgcgtgtgt	300
cagctgcacg	ggcacgttgt	tctcggtcag	gaagatgtac	agcttgcgttgc	cgtgtcgta	360
ctcctcggtc	gcgtggtcga	acaggaacag	gccccgcgc	tccaggctgt	gggtgttagca	420
caagctgttc	atgcgtatgt	acaggttgc	gtctgcata	tccgttgtca	cctgtcggtt	480
cagcaggttgc	atgtgtcg	ctcgggattt	tatctggaa	ctgttattca	gcgttcgtc	540
cctgtaaaca	tcgtgtatcat	aggtaccatt	tctaattgt	ccatgtcagg	cattatcaca	600
cttatgtat	atttaaagc	atccgttccc	catgtctct	gcattttctc	taagctgttt	660
cttcgttcc	tcaaaaagtt	tgttcatctc	tgaatcggt	aggtcaatag	tgtgtgtt	720
ctccagagcc	accagcagct	cggcattgt	ggaccacaga	tcaatcttgc	tatcctcaac	780
atacttctcc	aggctctgt	tgcggccctc	gacttcggag	aactcttctc	cgatctgt	840
gaatttttcg	ttgggtttcc	cgatcagtct	gttgagctt	ccattgtatct	gatctatagc	900
ggcctgcgtg	ctcttgcgt	cggcagcctg	gcctctgcct	tccggagttt	gatgtctaaa	960
accataccaa	ccatccacca	tccctccca	gccattctcg	ataaagctg	cgatagcccc	1020
aaatatcccg	cgggctgtt	tttggggac	attgcgcata	ccggtegc	gttcaagg	1080
attttgc	acgtaccgt	gacaggcccc	ataagttt	cgatcgt	tctggaaagg	1140
tttgcatttgc	ggaatgctcc	cattgggtgt	gatacactca	gagttgcatt	ttccgatcg	1200
tgcatactgt	ctcataatgg	aggacttcc	agagcggatt	ttaaagtatc	cccgggggc	1260
gatcagggttgc	ccagttactgt	taatcaggag	aatatcgct	ggttacg	tggtcagta	1320
aatactgtat	cgactaggaa	tgttccgtac	tctgggtct	ctgcccata	tgggagacac	1380
agtttgcgttgc	cttcttttgc	tagacacgt	aatcctgcct	gaagcctgg	catacagaaa	1440
gatctgtatcc	ttgtcagtgc	caggatgt	aactccccaa	atgtaaagct	tatcgaactg	1500
ctcggttgc	ggcatggta	cattcagtgc	agggtaacttgc	aagttcagat	gtgtcagcc	1560
atccgtctgc	ctgaagaaac	tattttgc	tctccttgc	caagcgg	acacgtcc	1620
ctgtgtaaacg	ccagttccgt	taaatgactc	gttgttgc	tccagagtcc	cactagacgc	1680
gacgaggggac	cgcaggctgg	cataatcagg	cacatcgaa	gggttagca	ttgaataagc	1740
ctttgaccc	tccacaaaca	gatcccactt	cttatttgc	aatccgtc	actgagggtc	1800
gcaccaacagc	gcgtcgatca	gcgtacaatt	cttccgtcc	agaatctgg	ggggactatc	1860
gcagatctct	cctgtggagc	tactttgcac	cagctcggt	gcattagtc	cctctatctg	1920
atcggttagt	atggtttca	caatggccc	gttagggac	gcatggtgc	ccaaaggcag	1980
cgtggctgtt	gagttgtcat	tgccggcag	tttctgggc	aacaccagac	acagttatgt	2040
ggacagcgca	attatggttt	tcat				2064

-continued

<210> SEQ ID NO 64
<211> LENGTH: 2058
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 64

atgaaagtga agctgctggc	60
gttatcggtt accacgcgg	120
gtgaccgtgt cccacagcgt	180
ctgaaaaggca ttgccccttc	240
aatcctgagt gcgagctgt	300
aaccctgaga atggcacctg	360
cagctgtcca gcgtgtccag	420
cccaatcata cagtgaccgg	480
tacagaaacc tgctgtggct	540
taacgccaaca acaaagaaaa	600
atcggcatcc agaaggccct	660
cactacagca gaaagttcac	720
ggcaggatca actactactg	780
aacggcaatc tgatcgcccc	840
atcatcaaca gcaacgcccc	900
gctatcaata gcagectgcc	960
aaatacgtgc ggagcgccaa	1020
cagagcagag gcctgtttgg	1080
gtggatgggt ggtacggcta	1140
cagaagagca cccagaacgc	1200
aagatgaaca cccagttcac	1260
gaaaacctga acaagaaggt	1320
ctcctggtcc tcctggaaaa	1380
ctgtacgaga aagtgaagag	1440
ttcgagttct accacaagtg	1500
gactacccca agtacagcga	1560
gacatcatca agctgctgaa	1620
atgagcatga gcagctgggt	1680
gaccacgccc ccgaggagta	1740
aacgtgcccc tgcaagtgac	1800
cagatcttcc agaaggccta	1860
gtggaccacg ccatcaagag	1920
ggcgcggcgc acgaggagga	1980
ggcaacgaga accacggcct	2040
aggaagagcg gatccctag	2058

-continued

<210> SEQ_ID NO 65
 <211> LENGTH: 685
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

 <400> SEQUENCE: 65

Met	Lys	Val	Lys	Leu	Leu	Val	Leu	Leu	Cys	Thr	Phe	Thr	Ala	Thr	Tyr
1				5			10						15		
Ala	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
	20					25							30		
Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
	35				40							45			
Leu	Leu	Glu	Asn	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile
	50					55						60			
Ala	Pro	Leu	Gln	Leu	Gly	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly
	65					70						75			80
Asn	Pro	Glu	Cys	Glu	Leu	Leu	Ile	Ser	Lys	Glu	Ser	Trp	Ser	Tyr	Ile
	85						90					95			
Val	Glu	Lys	Pro	Asn	Pro	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	His	Phe
	100						105					110			
Ala	Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe
	115					120						125			
Glu	Arg	Phe	Glu	Ile	Phe	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr
	130					135						140			
Val	Thr	Gly	Val	Ser	Ala	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe
	145					150						155			160
Tyr	Arg	Asn	Leu	Leu	Trp	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn
	165						170					175			
Leu	Ser	Lys	Ser	Tyr	Ala	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu
	180						185					190			
Trp	Gly	Val	His	His	Pro	Pro	Asn	Ile	Gly	Ile	Gln	Lys	Ala	Leu	Tyr
	195					200						205			
His	Thr	Glu	Asn	Ala	Tyr	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg
	210					215						220			
Lys	Phe	Thr	Pro	Glu	Ile	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Gln	Glu
	225					230						235			240
Gly	Arg	Ile	Asn	Tyr	Tyr	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile
	245						250					255			
Ile	Phe	Glu	Ala	Asn	Gly	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala
	260					265						270			
Leu	Ser	Arg	Gly	Phe	Gly	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met
	275					280						285			
Asp	Lys	Cys	Asp	Ala	Lys	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser
	290					295						300			
Ser	Leu	Pro	Phe	Gln	Asn	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro
	305					310						315			320
Lys	Tyr	Val	Arg	Ser	Ala	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn
	325						330					335			
Ile	Pro	Ser	Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe
	340						345					350			
Ile	Glu	Gly	Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His
	355						360					365			

-continued

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
 370 375 380

Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
 385 390 395 400

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile
 420 425 430

Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
 435 440 445

Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
 450 455 460

Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
 465 470 475 480

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
 485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
 500 505 510

Arg Glu Lys Ile Asp Ser Gly Gly Asp Ile Ile Lys Leu Leu Asn Glu
 515 520 525

Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met Ser
 530 535 540

Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe
 545 550 555 560

Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe
 565 570 575

Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro
 580 585 590

Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu
 595 600 605

His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala
 610 615 620

Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val
 625 630 635 640

Ala Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys
 645 650 655

Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln
 660 665 670

Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 675 680 685

<210> SEQ_ID NO 66

<211> LENGTH: 2058

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 66

ctaggatccg ctcttcctgc tcttggcgat gcccttcacg tactggtcgg ccaggtacag 60

gccgtggttc tcgttgccga tcagtcgat cttgtccagg atgtccttga acagcacctc 120

ctcctcgtgc tgctcggcca cgtaccactg caggaagtgg aaggtggcgt ggtccttgct 180

cttgatggcg tggtccacga tggtgttat gctctcgctg atgtgctgct cgtgctcgta 240

201

202

-continued

ggccttctgg aagatctggg tcaggccctc gaacttgc tcggggcgca tggatgttgt
cagctgcacg ggcacgttgt tctcgtaga gaagatgatc agcttcttgc cgtgtcgta
ctctctggcg ggcgtggcga acaggaaacag gcccggcccg tccaggtgtt ggggttagca
ccagctgttc atgctcatgt acagggttgt gctctgcattt cctgtcgat
cagcagcttgc atgatgtcgc ctccggaaatc gatcttctcc cggttcaatc tgctttcc
gctgtacttg gggtagtcgtt aggttgcgtt ctgcacgtt tccatgtactt cgtcggt
cttgcgttag aactcgaagc agccgttgcc gatcttttgc gctgttttct tcaactgg
cttcacttgc tgcgtacaggt tcttcacgtt gctgtcggtt aagtccagggg tcccttcatt
ttccaggagg accaggagggtt cggcggttgc ggtccagatc tgatgaaagc cgtcgtcc
cttcttgcgttcc aggttttcca tccggcggttcc cagttgttgc aactcttgc ccacggcg
gaactgggtt ttcatcttctt cgttacacgtt gttcacatttgc ttggatgtgc cggt
gttctgggtt ctcttcgtat cggcggttgc gccagatccc tgctcatttctt ggtgggt
ggcgtaaccac ccattccacca tgcctgttca tccggcccttgc ataaagccgg caatggcg
aaacaggccctt ctgcgttgcgat tgctggggat attcctcagg cgggttccatca ttctc
ggcgctccgc acgttatttag gacactcgcc gatggtcaca ggggtccat tctggaaagg
caggctgttca ttgtatgttgc cctgtgggttgc ctgtacacttgc ggttacactt tgc
ggcggttgcgtt tgatgtatgc cgctgcacaaa gcctctgttgc agggcaaaagg cgtatctt
ggcgatcaga ttggcggttgc cctcgttgc gatgggttgc ccagggttca gcagggttca
gttagttagtttgc atccctgcctt cctgttgcgtt cactttttgtt ctggggat tgc
gaactttcttgc ctgttagtggc tggacaccac ggacacgttgc gatcttgc tgc
ggccttctgg atgcccgtatgttgc taggggggttgc gtgcacttcc cagaggacca
gcacttcttgc ttctttgttgc ttggcgtagc tcttgcgttgc gttggggat tgc
ccacaggagg tttctgttgc agctgttgc tccatgttgc ctacaaggagg cgtt
ggtcaactgttgc tgattggggcc agctgttgc tttggggat tgc
ggacacgttgc gacacgttgc ctctcgttgc ctcgttgc gcaaggatggc cagggt
ggtgcatttgc tcaagggttgc gtttgcgttgc gatgttgc gaggacttgc tggaaatc
cagctcgatc tcaagggttgc ccagaatccatc gcccggccatc ctacaatttgc
agggggcaatc ctttcagca gacacgttgc gccgttgc tgc
gtgtgggttgc acgggttgc tcttttccatc ctcgttgc gcaaggatggc tggaaatc
ggcggttgc tcaagggttgc ccagaatccatc gccgttgc tgc
cagcagcttgc actttcatc

```
<210> SEQ ID NO 67
<211> LENGTH: 2112
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 67

atgaaggcca tcatcggtct gctgatggtg gtcacaagca acgcccatacg aatctgtacc	60
ggcatcaccca gcagcaatag ccctcactgcgtc gtggaaaacag ctacacaggg cgaagtgaat	120
gtgacccggcg tgatccctct gaccacaaca cctacaaaga gccacttgcg caatctgaag	180
ggcacagaga caagaggcaa gctgtgtccc aagtgcctga attgcacaga tctggatgtg	240

-continued

gctctggca gacctaagt tacaggcaa atccctagcg ccagagtgtc cattctgcat 300
 gaagtgcac ctgtgaccag cggctgttt cctattatgc acgaccggac caagatcaga 360
 cagctgccta atctgtgag aggctacgag cacatcagac tgagcaccca caatgtgatc 420
 aacgcccggaa atgctcctgg cggcccttat aagatcggca catctggcag ctgccccaaac 480
 attacaatg gcaatggctt ctttgccacc atggcttggg ccgtgctaa gaacgataag 540
 aacaagaccc ccaccaaccc cctgacaatc gaggtgccc atatctgtac agagggcag 600
 gatcagatca ccgtgtgggg atttcacagc gacaacgaaa cacagatggc caagctgtac 660
 ggcgatagca agcctcagaa gtttaccgc tctgccaatg gctgtgaccac acactatgt 720
 ttcagatcg gcgggttccc taatcagaca gaagatggcg gactgctca gtctggaaga 780
 atcgtgggttggg attacatggt gcagaagtct ggcaagaccc gcaccatcac atatcagaga 840
 ggaatctgc tgccccagaa agtgtggc gcttctggaa gatccaaagt gatcaagggc 900
 agcctgcctc tgattggaga agccgattgt ctgcacgaga aatacggcg cctgaacaag 960
 agcaagccctt actatacagg cgagcacgccc aaggccatcg gcaattgtcc tatttgggtc 1020
 aagacccctc tgaagtcgc caatggcaca aagtatagac ctccagccaa gctgtgaaa 1080
 gagagaggct ttttggage tatcgccggc tttctggaa goggatgggaa gggatgatt 1140
 gctggatggc atggcacac atctcatggc gcacatggc tggcagtggc tgctgatctg 1200
 aaatctacac aggaagccat caacaagatc accaagaacc tgaacagcct gagcgagctg 1260
 gaagtgaaga atctgcagag actgtctggc gccatggacg aactgcacaa tgagatcctg 1320
 gaactggacg agaagggtgga ctagtgcaga gcccataaca tcagcagccaa gattgaactg 1380
 gctgtgtgc tgtctaacga gggcatcatc aatagcggg acgaacatct gctggccctg 1440
 gaaagaaaagc tgaagaagat gctgggaccc agcgccgtgg aaatcgccaa tggatgctt 1500
 gagacaaagc acaagtgc aa ccagacctc ctggatagaa ttggccggg aacatttgat 1560
 gccggcaggt tttctctgcc caccttcgtt agcctgaata tcacatccgg aggccacatc 1620
 atcaagtcgc tgaacgagca ggtgaacaag gagatgcaga gcagcaacct gtatcgagc 1680
 atgagcagct ggtgetacac ccacagccctg gacggcgcgg gcctgttccct gttcgaccac 1740
 gccggcagg agtacgagca cgccaaaggat ctgatcatct tcctgaacga gaacaacgtg 1800
 cccgtgcagc tgaccagcat cagcgcccc gggcacaatg tcgaggccct gacccagatc 1860
 ttccagaagg cctacgagca cgagcagcac atcagcggaa gcatcaacaa catcggtac 1920
 cacgcccata agagcaagga ccacgcccacc ttcaacttcc tgcagtggta cgtggccgg 1980
 cagcacgagg aggagggtgct gttcaaggac atcctggaca agatcgagct gatcgccaaac 2040
 gagaaccacg gcctgtaccc ggccgaccag tacgtgaagg gcatcgccaa gagcagggaaag 2100
 agcggatcct ag 2112

<210> SEQ ID NO 68
 <211> LENGTH: 703
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 68

Met	Lys	Ala	Ile
Ile	Ile	Val	Leu
Met		Met	Val
		Val	Thr
		Ser	Asn
			Asp

1	5	10	15
---	---	----	----

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys

-continued

20	25	30
Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr		
35	40	45
Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu Thr		
50	55	60
Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp Val		
65	70	75
Ala Leu Gly Arg Pro Lys Cys Thr Gly Lys Ile Pro Ser Ala Arg Val		
85	90	95
Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile		
100	105	110
Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly		
115	120	125
Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu Asn		
130	135	140
Ala Pro Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro Asn		
145	150	155
Ile Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro		
165	170	175
Lys Asn Asp Lys Asn Lys Thr Ala Thr Asn Pro Leu Thr Ile Glu Val		
180	185	190
Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe		
195	200	205
His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser Lys		
210	215	220
Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val		
225	230	235
Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro		
245	250	255
Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys		
260	265	270
Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val		
275	280	285
Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu		
290	295	300
Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys		
305	310	315
Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys		
325	330	335
Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr		
340	345	350
Arg Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile		
355	360	365
Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His		
370	375	380
Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu		
385	390	395
Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser		
405	410	415
Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met		
420	425	430
Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp		
435	440	445

-continued

Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu
 450 455 460
 Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu
 465 470 475 480
 Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly
 485 490 495
 Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp
 500 505 510
 Arg Ile Ala Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro Thr
 515 520 525
 Phe Asp Ser Leu Asn Ile Thr Ser Gly Gly Asp Ile Ile Lys Leu Leu
 530 535 540
 Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser
 545 550 555 560
 Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe
 565 570 575
 Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile
 580 585 590
 Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser
 595 600 605
 Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala
 610 615 620
 Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp
 625 630 635 640
 His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp
 645 650 655
 Tyr Val Ala Glu Gln His Glu Glu Glu Val Leu Phe Lys Asp Ile Leu
 660 665 670
 Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala
 675 680 685
 Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 690 695 700

<210> SEQ_ID NO 69
 <211> LENGTH: 2112
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic
 <400> SEQUENCE: 69

ctaggatccg	ctttcctgc	tcttggcgat	gcccttcacg	tactggtcgg	ccaggtacag	60
gcccgtggttc	tcgttgccga	tcagctcgat	cttgtccagg	atgtccttga	acagcacctc	120
ctccctcggtgc	tgctcggcca	cgtaccactg	caggaagtgg	aagggtggcgt	ggtccttgc	180
cttgcgtggcg	tggtccacga	tgttgttgc	gctctcgctg	atgtgcgtgc	cgtgcgtcgt	240
ggcccttctgg	aagatctggg	tcaggccctc	gaacttgtgc	tggggggcgc	tgttgtgtt	300
cagctgcacg	ggcacgttgt	tctcggtcag	gaagatgtac	agcttctgg	cgtgcgtcgt	360
ctccctcgccg	gcgtggtcga	acaggaacag	gccggcgccc	tccaggctgt	gggtgttagca	420
ccagctgcgtc	atgctcatgt	acaggttgc	gctctgcac	tccttgcata	cctgcgtcgtt	480
cagcagcttg	atgatgtcgc	ctccggatgt	gatattcagg	ctatcgaagg	tgggcagaga	540
aaactcgccg	gcatcaaatg	ttccggcgcc	aattctatcc	aggcaggtct	ggttgcactt	600

209

210

-continued

<210> SEQ ID NO 70

<211> LENGTH: 837

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 70

atgaaggcca	agctgtgttgt	gctgtgtgc	accttacccg	ccacctacgc	cgacaccatc	60
tgcattggct	accacgccaa	caacagcacc	gacaccgtgg	ataccgtgct	ggaaaagaac	120
gtgaccgtga	cccacacgt	gaacctggga	tccggactga	aatggtcac	cggctgaga	180
aacatccccca	gcatccagag	cagaggcctg	tttggagcca	ttgcggctt	tattgaggc	240
ggatggacccg	aatggtgga	tgggtggtag	ggctaccacc	accagaatga	gcagggctct	300
ggctatggccg	ccgatccagaa	gtctaccacag	aacgccatca	acggcatcac	caacaaggta	360
aacagcgtga	tcgagaagat	gggcggcgat	cctgaatggg	acagagagat	caacaactac	420
accagcatca	tctacacgct	gatcgaggaa	agccagaacc	agcaggaaaa	cggcacaggc	480
ggcggatctg	gaattgtgca	gcagcagaac	aacctgtga	gagccattga	ggcccagcag	540
catctgtgc	agctgacagt	gtggggcatc	aagcagctgc	agacccatcaa	tgcccgagctg	600

-continued

```

ctggtcctcc tggaaaacga gagaaccctg gacttccacg acagcaacgt gaagaacctg      660
tacgagaaag tgaagtccca gctgaagaac aacgc当地aaag agatcggcaa cggctgcttc      720
gagttctacc acaagtgcaa caacgagtgc atggaaagcg tgaagaacgg cacctacgac      780
taccccaagt acagcgagga aagcaagctg aacagagaga agatcgactc cggaggc      837

```

<210> SEQ ID NO 71
<211> LENGTH: 279
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 71

```

Met Lys Ala Lys Leu Leu Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1           5           10          15

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
20          25           30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
35          40           45

Leu Gly Ser Gly Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser
50          55           60

Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
65          70           75           80

Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn
85          90           95

Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala
100         105          110

Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Gly
115         120          125

Gly Asp Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Ile Ile
130         135          140

Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn Gly Thr Gly
145         150          155          160

Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile
165         170          175

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
180         185          190

Leu Gln Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Asn Glu Arg
195         200          205

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
210         215          220

Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
225         230          235          240

Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys Asn
245         250          255

Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
260         265          270

Glu Lys Ile Asp Ser Gly Gly
275

```

<210> SEQ ID NO 72
<211> LENGTH: 837
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 72

-continued

```

gcctccggag tcgatttct ctctgttcag cttgcttcc tgcgtgtact tggggtagtc      60
gttaggtccg ttcttcacgc ttccatgca ctgcgttgtt cacttgtgtt agaactcgaa     120
gcagccgttg ccgatcttt tgccgttgtt cttcagctgg gacttcactt tctcgtacag     180
gttcttcacg ttgcgttgtt ggaagtccag ggttctctc tttccagga ggaccagcag     240
ctcggcatgt taggtctgca gctgcttgc gccccacact gtcagctgca gcagatgtg      300
ctggcgtca atggctctca gcagggttgtt ctgcgtgtgc acaattccag atccgccc      360
tgtgcgttt tcctgttgtt tctggcttc ctgcgtcagg ctgttagatga tgctggtgta     420
gttgggtatc tctctgtccc attcaggatc gcccgcacat ttctcgtatca cgctgttac     480
tttgggtgt atgcccgtga tggcggtctg ggtagacttc tgatcggccg catagccaga     540
gcctgtca ttctgttgtt ggtagccgtatcc accatcccg tccatccgc      600
ctcaataaaag cccgcaatgg ctccaaacag gcctctgttc tggatgttgg ggtatgttct     660
caggccgggtg accattctca gtccggatcc caggttcacg ctgtgggtca cggtcacgat     720
ctttccagc acggatcca cgggttgtt gctgttgttgc gcgtggtagc caatgcagat     780
ggtgtcggcg taggtggcg taaaggtgca cagcagcacc agcagcttgg ccttcat      837

```

<210> SEQ ID NO 73

<211> LENGTH: 828

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 73

```

atgaaggcta tcctgggtt gctgctgtac acctttgcca cccgcaatgc cgacaccctg      60
tgttattggct accacgccaa caacagcacc gacaccgtgg ataccgtgtt ggaaaagaac     120
gtgaccgtga cccacacgtt gaacctgggc tccggcgtga gactggccac cggcctgaga     180
aacatccccca gcattcagag cagaggcctg tttggagcca ttggcggctt tattgaggc     240
ggatggccgg gaatgggtt ggggttgttgc ggctaccacc accagaatga gcagggtct     300
ggctatgccg cccgacatcgaa gtctaccccg aacgcccattcg acgagatcac caacaagg     360
aacagcgtga tcgagaagat gggcggtctgg gacccatggg acagagagat caacaactac     420
accagcatca tctacagcct gatcgaggaa agccagaacc agcaggaaaa cggcacaggc     480
ggcggatctgta gatgggttgc gcacccatcgaa aacccatgtt gacccatgg ggcccgac     540
catctgtgc agctgacagt gtggggcatc aagcagctgc agacccatcaa cggccagctg     600
ctgggtgtgc tcgagaatga gagaaccctgg gactaccacg acagcaacgt gaagaacctg     660
tacggaaaag tgcggagcca gctgaagaac aacgccaag agatcggcaa cggctgttcc     720
gagttctacc acaagtgcga caataccctgc atggaaaaggc tgaagaacgg cacctacgac     780
taccccaagt acagcggatcgaa agccaaaggc aaccggaaag agatcgat      828

```

<210> SEQ ID NO 74

<211> LENGTH: 276

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 74

Met	Lys	Ala	Ile	Leu	Val	Val	Leu	Tyr	Thr	Phe	Ala	Thr	Ala	Asn
1														

5

10

15

Ala	Asp	Thr	Leu	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
20															

25

30

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

215

216

-continued

35	40	45	
Leu Gly Ser Gly Leu Arg	Leu Ala Thr Gly	Leu Arg Asn Ile Pro Ser	
50	55	60	
Ile Gln Ser Arg Gly	Leu Phe Gly Ala Ile	Ala Gly Phe Ile Glu Gly	
65	70	75	80
Gly Trp Thr Gly Met Val Asp	Gly Trp Tyr Gly Tyr His	His Gln Asn	
85	90	95	
Glu Gln Gly Ser Gly Tyr	Ala Ala Asp	Leu Lys Ser Thr Gln Asn Ala	
100	105	110	
Ile Asp Glu Ile Thr Asn Lys	Val Asn Ser Val	Ile Glu Lys Met Gly	
115	120	125	
Gly Trp Asp Pro Trp Asp Arg	Glu Ile Asn Asn Tyr	Thr Ser Ile Ile	
130	135	140	
Tyr Ser Leu Ile Glu Glu	Ser Gln Asn Gln	Glu Asn Gly Thr Gly	
145	150	155	160
Gly Gly Ser Gly Ile Val Gln Gln	Gln Asn Asn Leu	Leu Arg Ala Ile	
165	170	175	
Glu Ala Gln Gln His Leu Leu Gln	Leu Thr Val Trp	Gly Ile Lys Gln	
180	185	190	
Leu Gln Thr Tyr Asn Ala Glu	Leu Leu Val	Leu Glu Asn Glu Arg	
195	200	205	
Thr Leu Asp Tyr His Asp Ser Asn Val Lys	Asn Leu Tyr	Glu Lys Val	
210	215	220	
Arg Ser Gln Leu Lys Asn Asn Ala Lys	Glu Ile Gly Asn	Gly Cys Phe	
225	230	235	240
Glu Phe Tyr His Lys Cys Asp Asn Thr	Cys Met Glu Ser	Val Lys Asn	
245	250	255	
Gly Thr Tyr Asp Tyr Pro Lys Tyr	Ser Glu Glu Ala	Lys Leu Asn Arg	
260	265	270	
Glu Glu Ile Asp			
275			

<210> SEQ ID NO 75
<211> LENGTH: 828
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 75
atcgatctct tcccggttca gcttggcttc ctgcgtgtac ttggggtagt cgttaggtgcc 60
gttcttcacg ctttccatgc aggtattgtc gcacttggtagaactcga agcagccgtt 120
gcccgtatctt ttggcggttgt tcttcagctg gctccgcact ttctcgtaaa gggttcttcac 180
gttgctgtcg tggttagtcca gggttctctc attctcgagc agcaccagca gctcggcggtt 240
gttaggtctgc agctgcttga tgccccacac tgtcagatgc agcagatgtc gctgggcctc 300
aatggctctc agcagggttgt tctgctgctg cacaattcca gatccgcgc ctgtgccgtt 360
tttcgtgtgg ttctggctt cctcgatcg gctgttagatg atgctggtgt agttgttgat 420
ctctctgtcc catgggtccc agccgcacat cttctcgatc acgctgttca ctttgggttgt 480
gatctcgatcg atggcggttct gggtagactt caggtcgccgc gcatagccag agccctgtcc 540
attctgggtgg tggttagccgt accacccatc caccattccg gtccatccgc cctcaataaaa 600
gccggcaatg gtcacaaaca ggcctctgtc ctgaatgtcg gggatgttcc tcaggccgggt 660
qgccaatctc aqgcccqaqc ccqggttac qctgtqqqtc acqqtacqtc tctttccaq 720

-continued

cacggtatcc acgggtcggtcggtt ggcgtggtag ccaatacaca gggtgtcgcc	780
attggcggtg ccaaagggtgt acagcagcac caccaggata gccttcat	828

<210> SEQ ID NO 76
<211> LENGTH: 822
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 76

atggccatca tctacctgat tctgctgttt acagccgtca gaggcgatca gatctgtatt	60
ggctaccacg ccaacaatag caccgagaaa gtggataccca tcctggaaag aaatgtgaca	120
gtgacacacg ccaaggatat tggatcagga ctgggtgtgg ctacaggact gagaaatgtg	180
cctcagattg agagcagagg cctgtttgga gccatttgctg gctttattga aggccgatgg	240
cagggaatga ttgatgggtg gtacggctac caccactcta atgatcaggg atctggatat	300
gccggccaca aagaatctac acagaaagcc ttcgacggca tcaccaacaa agtgaatagc	360
gtgatcgaga agatgggcgg agatcccga tgggacagag agatcaacaa ctacaccagc	420
atcatctaca gcctgatcga ggaaagccag aatcagcagg aaaatggAAC aggccggagga	480
tctggaatttgcagcagca gaacaatctg ctgagagctta ttgaagctca gcagcatctg	540
ctgaatctga cagtgtgggg aatcaaacag ctgcagacat acaatgctga gctgctgg	600
ctgatggaaa atgagagaac cctggacttc cacgacagca atgtgaagaa cctgtacgac	660
aaagtgcgga tgcagctgag agacaatgtg aaagaactgg gcaatggctg cttcgagttc	720
taccacaagt ggcacgatga gtgtatgaac agcgtgaaga acggcaccta cgactaccct	780
aagtacgagg aagagagcaa gctgaacaga aatgagatca ag	822

<210> SEQ ID NO 77
<211> LENGTH: 274
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 77

Met Ala Ile Ile Tyr Leu Ile Leu Leu Phe Thr Ala Val Arg Gly Asp	
1 5 10 15	

Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Lys Val Asp	
20 25 30	

Thr Ile Leu Glu Arg Asn Val Thr Val Thr His Ala Lys Asp Ile Gly	
35 40 45	

Ser Gly Leu Val Leu Ala Thr Gly Leu Arg Asn Val Pro Gln Ile Glu	
50 55 60	

Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Trp	
65 70 75 80	

Gln Gly Met Ile Asp Gly Trp Tyr Gly Tyr His His Ser Asn Asp Gln	
85 90 95	

Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala Phe Asp	
100 105 110	

Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Gly Gly Asp	
115 120 125	

Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser	
130 135 140	

Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn Gly Thr Gly Gly	
145 150 155 160	

Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala

US 9,441,019 B2

219

220

-continued

165	170	175
Gln Gln His Leu Leu Asn Leu Thr Val Trp Gly Ile Lys Gln Leu Gln		
180	185	190
Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr Leu		
195	200	205
Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Met		
210	215	220
Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe		
225	230	235
Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly Thr		
245	250	255
Tyr Asp Tyr Pro Lys Tyr Glu Glu Ser Lys Leu Asn Arg Asn Glu		
260	265	270
Ile Lys		

<210> SEQ ID NO 78
<211> LENGTH: 822
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 78
cttgatctca tttctgttca gcttgctctc ttccctcgta ctttagggtagt cgtaggtgcc 60
gttcttcacg ctgttcatac actcatcgtc gcactttgtgg tagaactcga agcagaccatt 120
gcccagttct ttcacattgt ctctcagctg catccgcact ttgtcgtaa gggttcttcac 180
attgctgtcg tggaagtcca gggttctctc atttccatc agcaccagca gctcagcatt 240
gtatgtctgc agctgtttga ttccccacac tgcgcattc agcagatgct gctgagcttc 300
aatagctctc agcagattgt tctgtgtctc cacaattcca gatcctccgc ctgttccatt 360
ttcctgtctga ttctggctt cctcgatcg gctgttagatg atgctggtgt agttgttgc 420
ctctctgtcc cattcggat ctccgcctat cttctcgatc acgctattca ctttgggttgc 480
gatgcccgtcg aaggctttct gtgttagattc tttgtcgccg gcatatccag atccctgtatc 540
attagagtgg tggtagccgt accaccatc aatcatccca tgccatccgc cttcaataaaa 600
gccagcaatg gtcacaaaca ggcctctgct ctcaatctga ggcacatttc tcagtcctgt 660
agccagcacc agtcctgtatc caatatcctt ggctgtgttc actgtcacat ttctttccag 720
gtatggatcc actttctcggt tgctattgtt ggctgtgttag ccaatacaga tctgtatcgcc 780
tctgtacggatc gtaaacacca gaatcagatc gatgtatggcc at 822

<210> SEQ ID NO 79
<211> LENGTH: 858
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

```
<400> SEQUENCE: 79

atgaagacca tcatcgccct gagtacatc ttctgcctgg ccctgggcca ggacctgcc 60
ggcaacgaca acagcacccgc caccctgtgc ctggggccacc acggccgtgcc caacggcacc 120
ctggtaaga ccatcaccga cgaccagatc gaggtgacca acgccaccga gctgggctcc 180
ggcctgaagc tggccacccgg catgcggAAC gtgcccggAGA agcagacccg gggcctgttc 240
ggcgccatcg cgggtttcat cgagaacggc tgggaggggca tgatcgacgg ctggtaacggc 300
ttccggcacc agaacagcga gggcacccggc caggccggcc acctgaagag cacccaggcc 360
qccatcqacc aqatcaacqq caaqqtqaac cqqqtqatcq aqaaqaccqq cqqqqatccc 420
```

-continued

gagtgccacc	gggagatcaa	caactacacc	agcatcatct	acagcctgat	cgaggagac	480
cagaaccagc	aggagaacgg	caccggcgcc	ggcagcggca	tcgtgcagca	gcagaacaac	540
ctgctgcggg	ccatcgaggc	ccagcagcac	ctgctgcagc	tgaccgtgtg	gggcatcaag	600
cagctgcaga	gctacaacgc	cgagctgctg	gtggccctgg	agaaccagca	caccatcgac	660
ctgaccgaca	gagatgaa	caagctgttc	gagaagaccc	ggcggcagct	gcgggagaac	720
gcccggaca	tggcaacgg	ctgcttcaag	atctaccaca	agtgcgacaa	cgcctgcata	780
gagagcatcc	ggaacggcac	ctacgaccac	gacgtgtacc	gggacgaggc	cctgaacaac	840
cggttccaga	tcaaggcc					858

<210> SEQ ID NO 80

<211> LENGTH: 286

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 80

Met	Lys	Thr	Ile	Ile	Ala	Leu	Ser	Tyr	Ile	Phe	Cys	Leu	Ala	Leu	Gly
1									5			10			15

Gln	Asp	Leu	Pro	Gly	Asn	Asp	Asn	Ser	Thr	Ala	Thr	Leu	Cys	Leu	Gly
									20			25			30

His	His	Ala	Val	Pro	Asn	Gly	Thr	Leu	Val	Lys	Thr	Ile	Thr	Asp	Asp
									35			40			45

Gln	Ile	Glu	Val	Thr	Asn	Ala	Thr	Glu	Leu	Gly	Ser	Gly	Leu	Lys	Leu
									50			55			60

Ala	Thr	Gly	Met	Arg	Asn	Val	Pro	Glu	Lys	Gln	Thr	Arg	Gly	Leu	Phe
									65			70			80

Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Asn	Gly	Trp	Glu	Gly	Met	Ile	Asp
									85			90			95

Gly	Trp	Tyr	Gly	Phe	Arg	His	Gln	Asn	Ser	Glu	Gly	Thr	Gly	Gln	Ala
									100			105			110

Ala	Asp	Leu	Lys	Ser	Thr	Gln	Ala	Ala	Ile	Asp	Gln	Ile	Asn	Gly	Lys
									115			120			125

Leu	Asn	Arg	Val	Ile	Glu	Lys	Thr	Gly	Gly	Asp	Pro	Glu	Trp	Asp	Arg
									130			135			140

Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile	Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser
									145			150			160

Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr	Gly	Gly	Ser	Gly	Ile	Val	Gln	
									165			170			175

Gln	Gln	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	
									180			185			190

Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ser	Tyr	Asn	Ala	Glu
									195			200			205

Leu	Leu	Val	Ala	Leu	Glu	Asn	Gln	His	Thr	Ile	Asp	Leu	Thr	Asp	Ser
									210			215			220

Glu	Met	Asn	Lys	Leu	Phe	Glu	Lys	Thr	Arg	Arg	Gln	Leu	Arg	Glu	Asn
									225			230			240

Ala	Glu	Asp	Met	Gly	Asn	Gly	Cys	Phe	Lys	Ile	Tyr	His	Lys	Cys	Asp
									245			250			255

Asn	Ala	Cys	Ile	Glu	Ser	Ile	Arg	Asn	Gly	Thr	Tyr	Asp	His	Asp	Val
									260			265			270

Tyr	Arg	Asp	Glu	Ala	Leu	Asn	Asn	Arg	Phe	Gln	Ile	Lys	Gly		
									275			280			285

-continued

<210> SEQ ID NO 81
<211> LENGTH: 858
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 81

gcccttgatc	tggaaaccgg	tgttcagggc	ctcgccccgg	tacacgtcgt	ggtcgttaggt	60
gcccgttccgg	atgcgtctcg	tgcaggcgtt	gtcgcaactt	tggtagatct	tgaagcagcc	120
gttgcggatcg	tcctcgccgt	tctcccgag	ctgcccgg	gtcttcgt	acagcttgtt	180
catctcgctg	tccggcagg	cgtatgggtgt	ctgggttctcc	agggccacca	gcagctccgc	240
gtttagctc	tgcagctgt	tgtatgcccc	cacggtcag	tgcagcagg	gtctgtggc	300
ctcgatggcc	cgcagcagg	tgttctgt	ctgcacat	ccgctgccc	cgccggtgcc	360
gttctctgc	tgggtctgg	tctccctcgat	caggctgt	atgtatgtgg	tgttagttgtt	420
gatctcccg	tcccactcg	gatgcggcc	ggtcttcgt	atcaccgg	tcaagtggcc	480
gtttagtctgg	tccgtatggcgg	cctgggtgt	cttcagg	ggccgttgc	cggtgcctc	540
gctgttctgg	tgcggaga	cgttacc	gtcgatcg	ccctccca	cggttcgtat	600
gaagccggcg	atggcgccg	acaggcccc	ggtctgtt	tccggcact	tccgcattgc	660
ggtggccagg	ttcaggccgg	agccca	ggtggcgtt	gtcac	tctggctgt	720
ggttagtggc	ttcaccagg	tgccgttgg	cacggcgtt	tggcccagg	acagggtggc	780
ggtgctgtt	tgcgtggcgg	gcagg	tctgtt	gcccagg	aggcagaaga	840
ggcgatgtat	gtcttcat					858

<210> SEQ ID NO 82
<211> LENGTH: 867
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 82

atgaaaacca	tcattgcct	gagctacatc	ctgtgcctgg	tgttacacaca	gaagctgccc	60
ggcaacgata	ata	gacccgc	cacactgtgt	ctgggacacc	acgcccgtgc	120
atcgtgaaaa	caatcacaa	cgaccagatc	gaagtgacca	atgccacaga	gctggctcc	180
ggcctgaagc	tggccaccgg	catgagaaat	gtgcccaga	agcagaccag	aggcatctt	240
ggcgccattg	ccggcttat	cgagaatggc	tgggagg	tggatgg	tgttacggc	300
ttcagacacc	agaatagcga	ggaaatttgg	caggccgc	atctgaaatc	tacccaggcc	360
gecatcgacc	agatcaacgg	caagctgaa	aggctgatcg	gcaagaccgg	cgccgatccc	420
gagtgggacc	gggagatcaa	caactacacc	agcatcatct	acagcctgat	cgaggagac	480
cagaaccagg	aggagaacgg	cacccggcggc	ggcagcggca	tctgtcagca	gcagaacaac	540
ctgctgggg	ccatcgaggg	ccagcagcac	ctgctgcagc	tgaccgtgt	gggcatcaag	600
cagctcgaga	gctacaatgc	cgaactgt	gtcgccctgg	aaaaccagca	cacaattgtat	660
ctgacagaca	gtgagatgaa	taa	gtctgttc	gagaaaaacca	agaagcag	720
gccgaggaca	tggcaacgg	ctgcttcaag	atctaccaca	agtgcgacaa	cgcctgcac	780
ggcagcatca	gaaacggcac	ctacgaccac	gacgtgtaca	gagatgaggc	cctgaacaac	840
cggtttcaga	tcaagg	gtc	cgagg			867

<210> SEQ ID NO 83
<211> LENGTH: 289

-continued

<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 83

```

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Thr
1           5          10          15

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20          25          30

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
35          40          45

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Gly Ser Gly Leu Lys Leu
50          55          60

Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe
65          70          75          80

Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp
85          90          95

Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Ile Gly Gln Ala
100         105         110

Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys
115         120         125

Leu Asn Arg Leu Ile Gly Lys Thr Gly Gly Asp Pro Glu Trp Asp Arg
130         135         140

Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser
145         150         155         160

Gln Asn Gln Gln Glu Asn Gly Thr Gly Gly Ser Gly Ile Val Gln
165         170         175

Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu
180         185         190

Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ser Tyr Asn Ala Glu
195         200         205

Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp Ser
210         215         220

Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn
225         230         235         240

Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp
245         250         255

Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val
260         265         270

Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Ser Gly
275         280         285

```

Gly

<210> SEQ ID NO 84
<211> LENGTH: 867
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 84

```

gcctccggag cccttgatct gaaaaccgggtt gttcaggccc tcatcttgtt acacgtcgta      60
gtcgtaggtg ccgtttctga tgctgccat gcaggcggtt tcgcacttgtt ggttagatctt     120
gaaggcggcc ttgccccatgt cctcgccgtt ttctctcagc tgcttcttgg ttttctcgaa     180
cagcttatttc atctcaactgt ctgtcagatc aatttgtgtgc tgggtttcca gggcgaccag     240
cagttcggca ttgttagctct gcagctgctt gatgccccac acggtcagct gcagcaggta     300

```

-continued

ctgctggcc tcgatggccc gcagcagggtt gttctgtgc tgcacatgc cgctgccgc	360
gccccgtggcc ttcttcgtct gggttcgtct ctctcgatc aggctgtaga tgatgttgt	420
gtatgttgtt atctcccggt cccactcggtt atcgccgcgg gtcttgccga tcagccgtt	480
cagcttgcgg tttgatctggt cgtatggccgc ctgggttagat ttcaagatcggtt cggccgttcc	540
aattccctcg ctattctggt gtctgaagcc gtaccaccca tccaccatcc cctccagcc	600
attctcgata aagccggcaa tggcgccaaa gatgcctctg gtctgtttct cgggcacatt	660
tctcatgcgg gtggccagct tcaggccgga gcccagctct gtggcattgg tcacttcgat	720
ctggctgtt gtgattgtt tcacgtatgtt gccattagcc acggccgttgtt gtcccaagaca	780
cagtgtggcg gtgttattat cggtggccggg cagttctgtt gtgaacacca ggcacaggat	840
qtaqctcaqq qcaatqatqq ttttcat	867

<210> SEQ ID NO 85
<211> LENGTH: 837
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 85
atggagaaga tcgtgctgct gctggccatc gtgagccctgg tgaagagcga ccagatctgc 60
atcggttacc acgccaacaa cagcacccagg caggtggaca ccatcatggaa gaagaacctgt 120
accgtgaccc acgcccaggaa catcggtcc ggectgggtgc tggccaccgg cctgcggaaac 180
agccccccaggc gggagagccg gcggaaagaag cggggccctgt tcggggccat cgccggcttc 240
atcgaggggcg gctggcaggg catggtgacg ggctggtacg gttaccacca cagcaaccgag 300
cagggcagcg gctacgcccgc cgacaaggag agcacccaga aggccatcgaa cggcggtgacc 360
aacaagggtga acagcatcat cgacaagatg ggccggcgatc ccggatggga ccggggagatc 420
aacaactaca ccagcatcat ctacagcctg atcgaggaga gccagaacca gcaggagaac 480
ggcacccggcg gccggcagcg catcggtcag cagcagaaca acctgctgcg ggccatcgag 540
gcccgacgcg acctgctgca gctgaccctgt tggggcatca acgagctgca gacctacaac 600
ggcgagctgc tggtgctgat ggagaacggag cggaccctgg acttccacga cagcaacgtg 660
aagaacctgt acgacaaggat gcccgtcgat ctgcgggaca acgccaaggaa gctgggcaac 720
ggctgcttcg agttctacca caagtgcgac aacgagtgc tggagagcat ccggaaacggc 780
acctaataact accccccaggta cagcgaggag gccccggctgtaa acggggaggaa qatcaq 837

<210> SEQ ID NO 86
<211> LENGTH: 279
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 86

Arg-Glu-Ile-Ser-Ile-Gly-Tyr-His-Ala-Lys-Lys-Ser-Thr-Gly-Glu-His

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile

33 34 35 36 37 38 39 40 41 42 43 44 45
Gly Ser Gly Leu Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg

Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe

-continued

Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His
85 90 95

His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr
100 105 110

Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp
115 120 125

Lys Met Gly Gly Asp Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
130 135 140

Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn
145 150 155 160

Gly Thr Gly Gly Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu
165 170 175

Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly
180 185 190

Ile Lys Gln Leu Gln Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu
195 200 205

Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr
210 215 220

Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn
225 230 235 240

Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser
245 250 255

Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg
260 265 270

Leu Lys Arg Glu Glu Ile Ser
275

<210> SEQ ID NO 87
<211> LENGTH: 837
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 87

```

gtgtgatctcc tccccgttca gccccggcctc ctgcgtgtac tgggggttagt tgttagtgcc      60
gttccggatg ctctccatgc actcggtgtc gcacttgtgg tagaactcga agcagccgtt      120
gccccagttcc ttggcggtgt ccccgagctg cagccgcacc ttgtcgtaa gggtttcac      180
gttgctgtcg tggaagtcca gggccgcctc gttctccatc agcaccagca gctccggcgtt      240
gttaggtctgc agctgcttga tgccccacac ggtcagctgc agcagggtgt gctgggcctc      300
gatggccgcg acgagggtgt tctgctgtcg cacgatggcc ctgcccgcgc cggtgccgtt      360
ctcctgtgg ttctggctct cctcgatcg gctgttagatg atgctgggt gtgtgttgc      420
ctccccggtcc cactcggtat cgccgcggat cttgtcgatg atgctgttca ctttgtgggt      480
cacggccgtcg atggcccttc ggggtcttc cttgtcggtc gcgtagccgc tgccctgttc      540
gttgctgtgg tggttagccgt accagccgtc caccatgccc tgccagccgc cctcgatgaa      600
gccccggatg ggcggcaaca ggccccgtt cttccggccgg ctctcccgct gggggctgtt      660
ccggcaggccg gtggccagca ccaggccgg a cccgatgtcc tggcggtgg tcacggtcac      720
gttcttctcc atgatgggtt ccacctgtcc ggtgctgttg ttggcggtgt agccgatgca      780
gatctggtgc ctcttcacca ggctcacat ggccagcaggc agcacgatct tctccat      837

```

<210> SEQ ID NO 88
<211> LENGTH: 837
<212> TYPE: DNA

-continued

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 88

atgaaggcca	tcategtgct	gctgatggtg	gtgaccagca	acgcccata	aatctgcacc	60
ggcatcacca	gcagcaatag	cccccatgtg	gtaaaaacag	ccacccaggg	cgaagtgaat	120
gtgacaggcg	tgatccctc	gggatcagga	ctgaagctgg	ccaatggcac	caagtacaga	180
cctcccgcca	agctgctgaa	agagagaggc	ttcttggcg	ccattggcg	atttctggaa	240
ggcggctggg	agggaatgat	tgcggctgg	cacggctata	catctcatgg	ggcccatggc	300
gtggctgtgg	ccggcgatct	gaagtctacc	caggaagcca	tcaacaagat	caccaagaac	360
ctgaacagcc	tgagcgagct	ggaaggaggc	gaccccgagt	gggatcgca	aatcaacaac	420
tacacatcta	tcatctacag	tctgatttag	gaaagccaga	accagcagga	aatgggact	480
gggggaggct	ccggaatcgt	gcagcagcag	aacaatctgc	tgcgagccat	tgaagctcag	540
cagcacctgc	tgcagctgac	agtgtgggc	atcaagcagc	tgcaggggtc	ccagattgaa	600
ctggccgtgc	tgctgtccaa	cgagggcata	atcaacagcg	aggatgaaca	cctgctggcc	660
ctggaacgga	agctgaagaa	gatgctgggc	ccttctgccc	tggagatcgg	caacggctgc	720
ttcgagacaa	agcacaagt	caaccagacc	tgcctggata	aatcgccgc	tggcacctc	780
aatgcccggcg	agttcagcct	gcctaccttc	gacagcctga	atatcacctc	cgaggc	837

<210> SEQ ID NO 89

<211> LENGTH: 279

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 89

Met	Lys	Ala	Ile	Ile	Val	Leu	Leu	Met	Val	Val	Thr	Ser	Asn	Ala	Asp
1					5			10			15				
Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys
					20			25			30				
Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Gly
					35			40			45				
Ser	Gly	Leu	Lys	Leu	Ala	Asn	Gly	Thr	Lys	Tyr	Arg	Pro	Pro	Ala	Lys
					50			55			60				
Leu	Leu	Lys	Glu	Arg	Gly	Phe	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Leu	Glu
					65			70			75			80	
Gly	Gly	Trp	Glu	Gly	Met	Ile	Ala	Gly	Trp	His	Gly	Tyr	Thr	Ser	His
					85			90			95				
Gly	Ala	His	Gly	Val	Ala	Val	Ala	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Glu
					100			105			110				
Ala	Ile	Asn	Lys	Ile	Thr	Lys	Asn	Leu	Asn	Ser	Leu	Ser	Glu	Leu	Glu
					115			120			125				
Gly	Gly	Asp	Pro	Glu	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile
					130			135			140				
Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Glu	Asn	Gly	Thr	
					145			150			155			160	
Gly	Gly	Gly	Ser	Gly	Ile	Val	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	
					165			170			175				
Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys
					180			185			190				
Gln	Leu	Gln	Gly	Ser	Gln	Ile	Glu	Leu	Ala	Val	Leu	Leu	Ser	Asn	Glu
					195			200			205				

-continued

Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys
210 215 220

Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys
225 230 235 240

Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala
245 250 255

Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser
260 265 270

Leu Asn Ile Thr Ser Gly Gly
275

<210> SEQ ID NO 90

<211> LENGTH: 837

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 90

gcctccggag gtgatattca ggctgtcgaa gtaggcagg ctgaactcgc cggcattgaa	60
ggtgccagcg gcgattctat ccaggcaggt ctggttgcac ttgtgttttgc ttcgaagca	120
gcccgttgcgc atctccacgg cagaagggcc cagcatcttc ttcagettcc gttccagggc	180
cagcagggtgt tcatactcgc tggttatgtat gcccctcggtt gacagcagca cggccaggtc	240
aatctggggac ccctgcagct gcttgatgcc ccacactgtc agctgcagca ggtgtgtcg	300
agcttcaatg gctcgcagca gattgttctg ctgctgcacg attccggagc ctcccccaagt	360
cccattctcc tgctggttct ggctttccctc aatcagactg tagatgtatg atgtgttagtt	420
gttgatttcg cgatcccact cggggtcgc tccttccaggc tgcgtcaggc tggcagggtt	480
cttgggtatc ttgttatgg ctteccgggtt agacttcaga tggcggccca cagccacgccc	540
atggggccca tgagatgtat agccgtgcgc gccggcaatc attccctccc agccgccttc	600
cagaaaatccg gcaatggcgc caaagaagcc tctcttttca agcagcttgg cgggaggct	660
gtacttgggtt ccatggcgc gcttcagttc tgatcccaga gggatcacgc ctgtcacatt	720
cacttcggcc tgggtggctt ttttaccac atgggggctt ttgctgtgg tggccgggt	780
gcagattcta tcggcggttgc tggtcaccac catcagcagc acgtatgtgg cttcat	837

<210> SEQ ID NO 91

<211> LENGTH: 867

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 91

atgaaaaccca taatttgcgt gtcctacata ctgtgtctgg tggttgcaca gaaactggcg	60
ggcaatgaca actcaacagc cacgctctgc ttggggcacc atgcgtccc taacgggacc	120
atttgtaaaa ccattactaa cgatcagata gaggtgacta atgcaccgc gctggctcc	180
ggcttgaacat tggcggcccg tatgcgcata gtcccccggaaa aacagacccg cgggatattt	240
ggggctatcg caggctttat cgagaatggc tggaaaggga tggatggatgg ttggatgtgt	300
ttagacatc aaaactccga aggagagggc caggctgcgc atctcaagag cacgcaggcc	360
gctatacatc agatcaatgg aaagctcaac agactgtatcg ggaaaacccg cggcgtatccc	420
gagtgggacc gggagatcaa caactacacc agcatcatct acagcctgtat cgaggagagc	480
cagaaccaggc aggagaacgg cacggccgc ggcagcggca tcgtgcagca gcagaacaac	540
ctgctgcgggg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag	600

-continued

cagctgcagt cctacaatgc cgagctgctg gtggctctgg agaatcagca cactattgac	660
ctgaccgatt cagagatgaa caaactttt gagaagacga agaaggcagct tagagaaaaat	720
gcagaggaca tggggAACGG atgcttaaa atatatcata agtgtgataa tgccctgcac	780
ggatcaatta gaaatggcacatgtatcacatgtttaca gggacgaagc gctgaataac	840
aggttccaga taaaaggctc cggaggc	867

<210> SEQ ID NO 92

<211> LENGTH: 289

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 92

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Ala			
1	5	10	15

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly			
20	25	30	

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp			
35	40	45	

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Gly Ser Gly Leu Lys Leu			
50	55	60	

Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe			
65	70	75	80

Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp			
85	90	95	

Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala			
100	105	110	

Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys			
115	120	125	

Leu Asn Arg Leu Ile Gly Lys Thr Gly Gly Asp Pro Glu Trp Asp Arg			
130	135	140	

Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser			
145	150	155	160

Gln Asn Gln Gln Glu Asn Gly Thr Gly Gly Ser Gly Ile Val Gln			
165	170	175	

Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu			
180	185	190	

Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ser Tyr Asn Ala Glu			
195	200	205	

Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp Ser			
210	215	220	

Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn			
225	230	235	240

Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp			
245	250	255	

Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val			
260	265	270	

Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Ser Gly			
275	280	285	

Gly

<210> SEQ ID NO 93

<211> LENGTH: 867

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

-continued

<400> SEQUENCE: 93

```

gcctccggag ctttatct ggaacctgtt attcagcgct tcttcctgt aaacatcgta 60
atcataggta ccattctaa ttgatccgat gcaggcatta tcacacttat gatatattt 120
aaagcatccg ttccccatgt cctctgcatt ttctctaage tgcttctcg tcttctcaaa 180
aagtttgttc atctctgaat cggtcaggtc aatagtgtgc tgattctcca gagccaccag 240
cagctcggca ttgttaggact gcagctgctt gatgcccac acggtcagct gcagcaggta 300
ctgctgggcc tcgatggccc gcagcaggta gttctgtgc tgcacgatgc cgctggcc 360
gecggtgcgc ttctctgtct gggttctggct ctcttcgatc aggctgtaga tgatgtctgg 420
gtagttgttg atctcccggt cccactcggt atcgccgcgg gtttcccgta tcagtctgtt 480
gagcttcca ttgatctgat ctatagcgcc ctgcgtgttc ttgagatcggt cagccgtggcc 540
tctgccttcg gagtttgcgt gtctaaaacc ataccaacca tccaccatcc ctcccagcc 600
attctcgata aagcctcgca tagccccaaa tatcccgccg gtctgtttt cggggacatt 660
gegcataaccg gtcgcagtt tcaagccgga gcccagctcg gtggcattag tcacctctat 720
ctgatcgta gtaatggttt tcacaatggt cccgttaggg acggcatggt gccccaaagca 780
gagcgtggct gttgagttgt cattgccccgg cagttctgg gcaaaacacca gacacagtt 840
gtaggacagc gcaattatgg ttttcat 867

```

<210> SEQ ID NO 94

<211> LENGTH: 837

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 94

```

atgaaaagtga agctgctgggt gctgctgtgt accttaccc ccacctacgc cgataaccatc 60
tgtatcggtt accacgccaa caatagcacc gacaccgtgg ataccgtgtt ggaaaagaac 120
gtgaccgtga cccacagcgt gaacctggga tcaggactga gaatgggtac cggcctgagg 180
aatatccccca gcataccagag cagaggccctg tttggcgccta ttggccggctt tatcgaggcc 240
ggatggacag gcatgggtggaa tgggtggtaac ggctaccacc accagaatga gcagggatct 300
ggctatgcgc ccgatcagaa gageacccag aacgcccatac acggcatcac caacaaagt 360
aacagegtga tcgagaagat gggcggcgat cctgaatggg acagagagat caacaactac 420
accagcatca tctacagcct gatcgaggaa agccagaacc agcaggaaaa cggcacaggc 480
ggcggatctg gaattgtgca gcagcagaac aacctgtga gagccattga ggcccagcag 540
catctgtgc agctgacagt gtggggcatc aagcagctgc agacctacaa cggcgaactc 600
ctggtcctcc tggaaaatga gaggaccctg gacttccacg acagcaacgt gaagaactg 660
tacgagaaag tgaagagcca gctgaagaac aacgccaaag agatcggcaa cggctgcttc 720
gagttctacc acaagtgcac cgacgagtgc atggaaagcg tgaagaacgg cacctacgac 780
taccccaagt acagcggagga aagcaagctg aaccgggaga agatcgattc cggaggc 837

```

<210> SEQ ID NO 95

<211> LENGTH: 276

<212> TYPE: PRT

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 95

Met	Lys	Val	Lys	Leu	Leu	Val	Leu	Leu	Cys	Thr	Phe	Thr	Ala	Thr	Tyr
1				5			10			15					

-continued

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
 20 25 30
 Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
 35 40 45
 Leu Gly Ser Gly Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser
 50 55 60
 Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
 65 70 75 80
 Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn
 85 90 95
 Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala
 100 105 110
 Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Gly
 115 120 125
 Gly Asp Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Ile Ile
 130 135 140
 Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn Gly Thr Gly
 145 150 155 160
 Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile
 165 170 175
 Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
 180 185 190
 Leu Gln Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
 195 200 205
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 210 215 220
 Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 225 230 235 240
 Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn
 245 250 255
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 260 265 270
 Glu Lys Ile Asp
 275

<210> SEQ ID NO 96

<211> LENGTH: 837

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 96

```

gcctccggaa tcgatcttct cccggttcag ctggcttcc tcgtgtact tggggtagtc      60
gtagggtccg ttcttcacgc tttccatgca ctctcggttg cacttgttgtt agaactcgaa    120
gcagccgttg ccgatctctt tggcggttggt cttcagctgg ctcttcactt tctcgatcac     180
gttcttcacg ttgctgttgtt ggaagtccag ggtccctctca ttttccaggaa ggaccaggag    240
ttcggcggttg taggtctgca gctgcttgat gccccacact gtcagctgca gcagatgtg      300
ctggggcctca atggctctca gcagggttgtt ctgctgctgc acaattccag atccggccccc    360
tgtgccgttt tcctgcttgtt tctggcttcc ctgcgttcagg ctgttagatga tgctgggtta    420
gttgttgcac tctctgtccc attcaggatc gcccggccatc ttctcgatca cgctgttcc      480
tttgttggtg atgcgggtga tggcggttctg ggtgctttc tgatccggcatagccaga      540
tccctgctca ttctgggtgtt ggtagccgta ccacccatcc accatgcctg tccatccggcc    600

```

-continued

```

ctcgataaaag ccggcaatgg cgccaaacag gcctctgctc tggatgctgg ggatattcct    660
caggccggtc accattctca gtcctgtatcc caggttacag ctgtgggtca cggtcacgtt    720
ctttccagc acggtatcca cggtgtcggt gctattgttg gctgtggtagc cgatacagat    780
ggtatacgcc taggtggcggt taaaggtaca cagcagcacc agcagctca ctttcat    837

```

<210> SEQ ID NO 97
<211> LENGTH: 837
<212> TYPE: DNA
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 97

```

atgaaggcca tcatcggtct gctgtatggc gtcacaagca acgcccata aatctgtacc    60
ggcatcacca gcagcaatacg ccctcacgtc gtgaaaacag ctacacaggg cgaagtgaat    120
gtgaccggcg tgatccctct gggatcagga ctgaagctgg ccaatggcac aaagtataga    180
cctccagccca agctgtgaa agagagaggc ttttttggag ctatcgccgg ctttctggaa    240
ggcggatggg agggaatgat tgctggatgg catggctaca catctcatgg cgcacatggc    300
gtggcagtgg ctgtgtatct gaaatctaca caggaagccca tcaacaagat caccaagaac    360
ctgaacagcc tgagcgagct ggaaggaggc gaccccgagt gggatcgccgaa atcaacaac    420
tacacatcta tcatactacag tctgtattgg gaaagccaga accagcagga gaatggact    480
ggggggaggct ccggaaatcggt gcagcagcag aacaatctgc tgcgagccat tgaagctcag    540
cagcacatgc tgcaatcgac agtgtggggc atcaagcagc tgcagggggccagattgaa    600
ctggctgtgc tgctgtctaa cgagggcattc atcaatagcg aggacgaaca tctgtggcc    660
ctggaaagaa agctgaagaa gatgtgggaa cctagcgcggc tggaaatccg caatggatgc    720
ttttagacaa agcacaagtg caaccagacc tgcctggata gaattgccgc cggAACATT    780
gatgcggcg agttttctct gcccaccttc gatagcctga atatcacatc cggaggc    837

```

<210> SEQ ID NO 98
<211> LENGTH: 279
<212> TYPE: PRT
<213> ORGANISM: Influenza virus

<400> SEQUENCE: 98

```

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1           5          10          15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20          25          30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Gly
35          40          45

Ser Gly Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys
50          55          60

Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu
65          70          75          80

Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His
85          90          95

Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu
100         105         110

Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu
115         120         125

Gly Gly Asp Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Ile
130         135         140

```

-continued

Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Asn Gly Thr
 145 150 155 160

Gly Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala
 165 170 175

Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys
 180 185 190

Gln Leu Gln Gly Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu
 195 200 205

Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys
 210 215 220

Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys
 225 230 235 240

Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala
 245 250 255

Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser
 260 265 270

Leu Asn Ile Thr Ser Gly Gly
 275

<210> SEQ ID NO 99

<211> LENGTH: 837

<212> TYPE: DNA

<213> ORGANISM: Influenza virus

<400> SEQUENCE: 99

gcctccggat gtgatattca ggctatcgaa ggtggcaga gaaaactcgc cggcatcaaa 60

tgttccggcg gcaattctat ccaggcaggt ctgggtgcac ttgtgttttg tctcaaagca 120

tccatttgcg atttccacgg cgtaggtcc cagcatcttc ttcatgtttc tttccaggc 180

cagcagatgt tcgttctcg tattgtat gcccgttta gacagcagca cagccagttc 240

aatctggctc ccctgcagct gcttgatgcc ccacactgtc agctgcagca ggtgtgtc 300

agcttcaatg gtcgcagca gattgttctg ctgctgcacg attccggagc ctccccagt 360

cccatttctcc tgctggttct ggcttccttc aatcagactg tagatgtatag atgtgttagtt 420

gttgatttcg cgatccact cggggtcgccc tccttcaggc tcgctcaggc tggtcagggt 480

cttgggtatcc ttgttgcgtt agatccatgc tcagcagcca ctgccacgccc 540

atgtgcgcaca tgagatgtgt agccatgcaca tccagcaatc attccctccc atccgccttc 600

cagaaagccg gcgatagctc caaaaaagcc tctcttttc agcagttgg ctggaggct 660

atactttgttgcatggccca gcttcagttcc tgatcccaga gggatcacgc cggtcacatt 720

cacttcgcggcc ttttgcgttgc ttttgcgttgc tgatgggttttttgcgttgc 780

acagattcta tcggcggttgc ttgtgaccac catcagcagc acgtatgtatgg cttcat 837

<210> SEQ ID NO 100

<211> LENGTH: 1332

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 100

atgaaggccca agctgtgttgc acctttaccc ccacacctacgc cgacaccatc 60

tgcattggctt accacgccaa caacagcacc gacaccgtgg ataccgtgttgc ggaaaagaac 120

gtgaccgttgc cccacacgttgc gaaacctggca tccggacttgc gaatgggtcac cggcctgttgc 180

-continued

aacatccccca	gcatccagag	cagaggcctg	tttggagcca	ttgccggctt	tattgagggc	240
ggatggccg	gaatgggttga	tggttgttac	ggctaccacc	accagaatga	gcaggcctct	300
ggctatgccg	ccgatcagaa	gtctacccag	aacgccatca	acggcatcac	caacaaagtg	360
aacagcgtga	tcgagaagat	gggcggcgt	cctgaatggg	acagagagat	caacaactac	420
accagcatca	tctacagcct	gatcgaggaa	agccagaacc	agcagggaaaa	cggcacaggc	480
ggcggatctg	gaatttgca	gcagcagaac	aacctgtga	gagccattga	ggcccagcag	540
catctgtgc	agctgacagt	gtggggcata	aagcagctgc	agacctacaa	tgccgagctg	600
ctgggtctcc	tggaaaacga	gagaaccctg	gacttccacg	acagcaacgt	gaagaacctg	660
tacgagaaaag	tgaagtccca	gctgaagaac	aacgccaaag	agatcggcaa	cggctgttc	720
gagttctacc	acaagtgc当地	caacgagtc当地	atggaaagcg	tgaagaacgg	cacctacgac	780
taccccaagt	acagcggagga	aagcaagctg	aacagagaga	agatcgactc	cggaggcgac	840
atcatcaagc	tgctgaacga	gcagggtgaac	aggagatgc当地	agagcagc当地	cctgtacatg	900
agcatgagca	gctgggtgcta	cacccacacgc	ctggacggcg	cggcctgtt	cctgttcac	960
cacggcccg	aggagtagca	gcacgccaag	aagctgtatca	tcttcctgaa	cgagaacaaac	1020
gtgcccgtgc	agctgaccag	catcagcgc当地	cccgagcaca	agttcgaggg	cctgaccag	1080
atcttcaga	aggcctacga	gcacgagcag	cacatcagcg	agagcatcaa	caacatcgtg	1140
gaccacgcca	tcaagagcaa	ggaccacgccc	accttcaact	tctgcagtg	gtacgtggcc	1200
gagcagcaccg	aggaggaggt	gctgttcaag	gacatcctgg	acaagatcga	gctgatcgcc	1260
aacgagaacc	acggcctgtta	cctggccgac	cagtacgtga	agggcatcgc	caagagcagg	1320
aagagcggat	cc					1332

<210> SEQ ID NO 101

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 101

Met	Lys	Ala	Lys	Leu	Leu	Val	Leu	Leu	Cys	Thr	Phe	Thr	Ala	Thr	Tyr
1								10							15

Ala	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
									20				25		30

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
									35			40		45	

Leu	Gly	Ser	Gly	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser
									50			55		60	

Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly
									65			70		75	80

Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn
									85			90		95	

Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	Asn	Ala
									100			105		110	

Ile	Asn	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Gly
									115			120		125	

Gly	Asp	Pro	Glu	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile	Ile
									130			135		140	

Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr	Gly
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

-continued

145	150	155	160
Gly	Gly	Ser	Gly
Ile	Val	Gln	Gln
165	170	175	
Asn	Asn	Leu	Leu
Arg		Ala	Ile
Glu	Ala	Gln	Gln
His	Leu	Leu	Gln
180	185	190	
Leu	Gln	Thr	Tyr
Asn	Ala	Glu	Leu
195	200	205	
Leu	Glu	Asn	Glu
Arg			
Thr	Leu	Asp	Phe
His	Asp	Ser	Asn
210	215	220	
Val	Lys	Asn	Leu
Tyr			Tyr
Gly			Glu
Cys			Lys
225	230	235	240
Asn	Asn	Ala	Ile
Glu			Gly
Asn			Asn
245	250	255	
Glu	Phe	Tyr	His
Asn	Cys	Asn	Asn
260	265	270	
Gly	Thr	Tyr	Asp
Tyr	Pro	Lys	Tyr
Ser	Glu	Glu	Ser
275	280	285	
Leu	Asn	Asp	Ser
Gly			Gly
Ile			Asp
290	295	300	
Leu	Lys	Leu	Met
Asn	Glu	Asn	Gln
305	310	315	320
Ser	Ser	Asn	Tyr
Leu	Tyr	Asp	Gly
Asp			Ala
325	330	335	
Gly			Gly
Ile			Asn
340	345	350	
Asn	Glu	Asn	Val
355	360	365	
Asp	Pro	Val	Gln
370	375	380	
Thr	Ser	Ile	Ser
385	390	395	400
Asn	Asp	Asp	His
395	400	405	
His	Ala	Thr	Phe
405	410	415	
Glu	Gln	Glu	Glu
420	425	430	
Val	Lys	Ile	Ala
435	440		

<210> SEQ ID NO 102
 <211> LENGTH: 1332
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 102

ggatccgctc	tccctgtct	tggcgatgcc	cttcacgtac	tggtcggcca	ggtacaggcc	60
gtggttctcg	tgcgcgatca	gctcgatctt	gtccaggatg	tccttgaaca	gcaccccttc	120
ctcgtgtc	tccgcacgt	accactgcag	gaagttgaag	gtggcggtgt	ccttgcttt	180
gatggcggtgg	tccacgtatgt	tgttgatgt	ctcgctgtatg	tgctgtcg	gctcgtaggc	240
cttctggaaag	atctgggtca	ggccctcgaa	cttgtcg	ggggcgctga	tgctggtcag	300
ctgcacgggc	acgttgttct	cgttcaggaa	gatgatcagc	ttcttggcggt	gctcgtaactc	360
ctcggcgccg	tggtcgaaca	ggaacaggcc	ggcccggtcc	aggctgtggg	tgttagcacca	420

-continued

gctgctcatg	ctcatgtaca	ggttgctgct	ctgcacatcc	ttgttcacct	gctcgttcag	480
cagcttgcgt	atgtcgccctc	cggagtcgt	cttctctctg	ttcagcttgc	tttccctcgct	540
gtacttgggg	tagtcgtagg	tgccgttctt	cacgcttcc	atgcactcgt	tgttgactt	600
gtggtagaaac	tcgaaggcage	cgttgcgcgt	ctctttggcg	ttgttcttca	gctgggactt	660
cacttctcg	tacaggttct	tcacggttct	gtcgttgcgt	tccagggttc	tctcgtttc	720
caggaggacc	agcagctcg	cattgttaggt	ctgcagctgc	ttgatgcccc	acactgtcag	780
ctgcagcaga	tgctgctggg	cctcaatggc	tctcagcagg	ttgttctgt	gctgacaaat	840
tccagatccg	ccgcctgtgc	cgttttcgt	ctgggttctgg	ctttccctcg	tcaggctgta	900
gatgatgctg	gtgttagttgt	tgtatctctct	gttccattca	ggatcgccgc	ccatcttctc	960
gatcacgctg	ttcaccttgc	ttgtgtatgcc	gttgtatggcg	ttctgggttag	acttctgtatc	1020
ggcggcatag	ccagagccct	gctcattctg	gtgggtggtag	ccgttaccacc	catccaccat	1080
tccggccat	ccgcctcaa	taaaggccgc	aatggctcca	aacaggccctc	tgcgttggat	1140
gttggggatg	tttctcaggc	cggtgaccat	tctcagtcgg	gatcccagg	tcacgctgtg	1200
ggtcacggc	acgttctttt	ccagcacgg	atccacgg	tccgtgtgt	tgttggcggt	1260
gttagccaatg	cagatgggt	cgccgttaggt	ggccgttaaag	gtgcacagca	gcaccaggag	1320
cttggcccttc	at					1332

<210> SEQ ID NO 103

<211> LENGTH: 1332

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 103

atgaaggcta	tcctgggtgt	gctgctgtac	acctttgc	ccgccaatgc	cgacaccctg	60
tgtattggct	accacgcca	caacagcacc	gacaccgtg	ataccgtgt	ggaaaagaac	120
gtgaccgtga	cccacagcgt	gaacctgggc	tccggcctga	gactggccac	ccgcctgaga	180
aacatccccca	gcattcagag	cagaggcctg	tttggagcca	ttggccggctt	tattgaggc	240
ggatggaccg	gaatggtgg	tgggtggatc	ggcttaccacc	accagaatga	gcagggtct	300
ggctatgccc	ccgacccatg	gtctacccat	aacgcccac	acgagatcac	caacaagt	360
aacagcgtga	tgcagaagat	ggccggctgg	gacccatggg	acagagat	caacaactac	420
accagcatca	tctacagcct	gatcgaggaa	agccagaacc	agcaggaaaa	cggcacaggc	480
ggcggatctg	gaatttgca	gcagcagaac	aacctgtca	gagccattga	ggcccagcag	540
catctgctgc	agctgacagt	gtggggcata	aagcagctgc	agacctacaa	cgccgagct	600
ctgggtgtgc	tgcagaatga	gagaaccctg	gactaccac	acagcaacgt	gaagaacctg	660
tacgagaaag	tgcggagcca	gctgaagaac	aacgccaag	agatcgccaa	cggtgtctc	720
gagttctacc	acaagtgcga	caatacctgc	atggaaagcg	tgaagaacgg	cacctacgac	780
taccccaagt	acagcggagga	agccaagctg	aaccgggaag	agatcgattc	cgaggccgac	840
atcatcaagc	tgctgaacga	gcaggtgaac	aaggagatgc	agagcagcaa	cctgtacatg	900
agcatgagca	gctgggtgta	cacccacacg	ctggacggc	ccggcctgtt	cctgttcac	960
cacgcccggc	aggagtacga	gcacgccaag	aagctgtatca	tcttcctgaa	cgagaacaac	1020
gtgcccgtgc	agctgaccag	catcagcgcc	cccgagcaca	agttcgaggg	cctgaccag	1080

-continued

atcttcaga	aggcctacga	gcacgagcg	cacatcagcg	agagcatcaa	caacatcg	1140
gaccacgcca	tcaagagcaa	ggaccacgcc	accttcaact	tcctgcagt	gtacgtggcc	1200
gagcagc	aggaggaggt	gctgttcaag	gacatcctgg	acaagatcga	gctgatcg	1260
aacgagaacc	acggcctgta	cctggccgac	cagtacgtga	agggcatcgc	caagagcagg	1320
aagagcggat	cc					1332

<210> SEQ ID NO 104

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 104

Met	Lys	Ala	Ile	Leu	Val	Val	Leu	Tyr	Thr	Phe	Ala	Thr	Ala	Asn
1				5			10						15	

Ala	Asp	Thr	Leu	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
			20				25						30		

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
			35			40						45			

Leu	Gly	Ser	Gly	Leu	Arg	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser
			50			55					60				

Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly
65				70			75					80			

Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn
			85				90					95			

Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Asn	Ala
			100			105					110				

Ile	Asp	Glu	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Gly
			115			120					125				

Gly	Trp	Asp	Pro	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile	Ile
			130			135					140				

Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr	Gly
145				150			155				160				

Gly	Gly	Ser	Gly	Ile	Val	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	
			165			170					175				

Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln
			180			185					190				

Leu	Gln	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu	Arg
			195			200					205				

Thr	Leu	Asp	Tyr	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys	Val
			210			215					220				

Arg	Ser	Gln	Leu	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys	Phe
225				230			235				240				

Glu	Phe	Tyr	His	Lys	Cys	Asp	Asn	Thr	Cys	Met	Glu	Ser	Val	Lys	Asn
			245			250					255				

Gly	Thr	Tyr	Asp	Tyr	Pro	Lys	Tyr	Ser	Glu	Glu	Ala	Lys	Leu	Asn	Arg
			260			265					270				

Glu	Glu	Ile	Asp	Ser	Gly	Gly	Asp	Ile	Ile	Lys	Leu	Asn	Glu	Gln	
			275			280					285				

Val	Asn	Lys	Glu	Met	Gln	Ser	Ser	Asn	Leu	Tyr	Met	Ser	Met	Ser	Ser
			290			295					300				

Trp	Cys	Tyr	Thr	His	Ser	Leu	Asp	Gly	Ala	Gly	Leu	Phe	Leu	Phe	Asp
305				310			315				320				

-continued

His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe Leu
325 330 335

Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu
340 345 350

His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His
355 360 365

Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile
370 375 380

Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala
385 390 395 400

Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile
405 410 415

Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr
420 425 430

Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
435 440

<210> SEQ ID NO 105

<211> LENGTH: 1332

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 105

ggatccgctc ttccctgtct tggcgatgcc ctgcacgtac tggtcggcca ggtacaggcc	60
gtgggttctcg ttgcgcgatca gctcgatctt gtccaggatg tccttgaaca gcacctcctc	120
ctcgtgtgtc tcggccacgt accactgcag gaagttaaga gttggcgttgt cttgtcttt	180
gtatggcgatgg tccacgatgt tggtgatgt ctgcgtatgt tgctgtctgt gtcgttaggc	240
cttctggaaat atctgggtca ggccctcgaa cttgtgtcg ggggcgtgta tgctggtcag	300
ctgcacgggc acgttggtct cggtcaggaa gatgatcagc ttcttggcggt gtcgtactc	360
ctcgccggcg tggtcgaaca ggaacaggcc ggcgcgtcc aggctgtggg tggtagcacca	420
gtctgtcatg ctcatgtaca gggtgctgt ctgcatctcc ttgttccacct gtcgttca	480
cacgttgcgtat atgtcgccctc cggaatcgat ctcttcccggt ttcaatggg cttcctcgct	540
gtacttgggg tagtcgttagg tgccgttctt caccgttccat atgcaggatgat tgctgcactt	600
gtggtagaaat tcgaaggcgc cggtgcgtat ctctttgggt ttgttctca gtcggctccg	660
cacatttcgtc tacagggtct tcacgttgcgt gtcgtggtag tccagggttc tctcatttc	720
gagcagcacc agcagctcg cggttaggt ctgcagctgc ttgtatgcacc acactgtcag	780
ctgcagcaga tgctgtgggg cctcaatggc tctcagcagg ttgttctgt gtcgcacaat	840
tccagatccg ccgcctgtgc cggttctgg ctgggttctgg ctttctcgat tcaggctgt	900
gatgatgtcg gtgtatgttgc tgatctcttc gtccatgggg tcccagccgc ccatcttc	960
gatcacgtcg ttcaatgttgc tggtgatctc gtcgtggcg ttctgggttag acttcagggtc	1020
ggcgccatag ccagagccct gtcattctcg gtgggttagt cggatccacc catccaccat	1080
tccggccat ccgcctcaa taaaaggccgc aatggctcca aacaggccctc tgctgtaat	1140
gctggggatg tttctcaggc cggtggccag tctcaggccg gagcccgatgt tcacgtgt	1200
ggtcacggtc acgttctttt ccagcacggat atccacgggt tcgggtgtgt tggtggcg	1260
gtagccaata cacagggtgt cggttgc ggtggcaaaag gtgtacagca gcaccaccag	1320

255

256

-continued

gatagccttc

1332

```
<210> SEQ ID NO 106  
<211> LENGTH: 1326  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:
```

<400> SEQUENCE: 106

atggccatca tctacctgat tctgtgttt acagccgtca gaggcgatca gatctgtatt
ggctaccacg ccaacaatag cacccgagaaa gtggatacca tcctggaaag aaatgtgaca 120
gtgacacacg ccaaggatata tggatcgaga ctgggtgtgg ctacaggact gagaaatgtg 180
cctcagattt agagcagagg cctgtttgga gccattgtg gctttattga aggccgatgg 240
cagggaatga ttgatgggtt gtacggctac caccactcta atgatcaggg atctggatata 300
gccggccgaca aagaatctac acagaaagcc ttgcacggca tcaccaacaa agtgaatagc 360
gtgatcgaga agatgggcgg agatcccga tgggacagag agatcaacaa ctacaccagc 420
atcatctaca gcctgatcga ggaaagccag aatcagcagg aaaatggAAC aggccgagga 480
tctggaaattt tgcaacatcg gaacaatctg ctgagagacta ttgaagctca gcagcatctg 540
ctgaatctga cagtgtgggg aatcaaacag ctgcagacat acaatgtga gctgtgggt 600
ctgatggaaa atgagagaac cctggacttc cacgcacagca atgtgaagaa cctgtacgac 660
aaagtgcgga tgcagctgag agacaatgtg aaagaactgg gcaatggctg cttcgagttc 720
taccacaagt ggcacgatga gtgtatgaac agcgtgaaga acggcaccta cgactaccct 780
aaagtacgagg aagagagcaa gctgaacaga aatgagatca agtccggagg cgacatcatc 840
aagctgtga acgagcaggt gaacaaggag atgcagagca gcaacctgtta catgagcatg 900
agcagctggt gctacacccca cagcctggac ggccgcggcc tggccctgtt cgaccacgccc 960
gccgaggaggat acgagcacgc caagaagctg atcatcttcc tgaacgagaa caacgtgccc 1020
gtgcagctga ccagcatcag cggccccggag cacaaggatcg agggcctgac ccagatcttc 1080
cagaaggccat acgagcacga gcagcacatc agcgagagca tcaacaacat cgtggaccac 1140
gccatcaaga gcaaggacca cgccaccccttc aacttccctgc aatgggtacgt ggccgagcag 1200
cacgaggagg aggtgctgtt caaggacatc ctggacaaga tcgagctgat cggcaacgag 1260
aaccacggcc tgcacccatcgcc cgaccagatc gtgaagggca tcgccaagag caggaagagc 1320
ggatcc

```
<210> SEQ ID NO 107
<211> LENGTH: 442
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 107

Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Lys	Val	Asp
20								25					30		

Thr Ile Leu Glu Arg Asn Val Thr Val Thr His Ala Lys Asp Ile Gly
 35 40 45

 Ser Gly Leu Val Leu Ala Thr Gly Leu Arg Asn Val Pro Gln Ile Glu
 50 55 60

-continued

Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp
 65 70 75 80

Gln Gly Met Ile Asp Gly Trp Tyr Gly Tyr His His Ser Asn Asp Gln
 85 90 95

Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala Phe Asp
 100 105 110

Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Gly Gly Asp
 115 120 125

Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser
 130 135 140

Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn Gly Thr Gly Gly
 145 150 155 160

Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala
 165 170 175

Gln Gln His Leu Leu Asn Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
 180 185 190

Thr Tyr Asn Ala Glu Leu Leu Val Met Glu Asn Glu Arg Thr Leu
 195 200 205

Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Met
 210 215 220

Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe
 225 230 235 240

Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly Thr
 245 250 255

Tyr Asp Tyr Pro Lys Tyr Glu Glu Ser Lys Leu Asn Arg Asn Glu
 260 265 270

Ile Lys Ser Gly Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn
 275 280 285

Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys
 290 295 300

Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala
 305 310 315 320

Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu
 325 330 335

Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys
 340 345 350

Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln
 355 360 365

His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser
 370 375 380

Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln
 385 390 395 400

His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu
 405 410 415

Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys
 420 425 430

Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 435 440

<210> SEQ ID NO 108
 <211> LENGTH: 1326
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 108

ggatccgctc ttccctgtct	tggcgatgc cttcacgtac	tggtcgccca ggtacaggcc	60
gtgggttctcg tggcgatca	gctcgatctt gtccaggatg	tccttgaaca gcaccccttc	120
ctcgtgtgc tcggccacgt	accactgcag gaagttgaag	gtggcggtgt ccttgcgttt	180
gatgggttgg tccacgtgt	tgttgcgtct ctcgtgtatg	tgctgtcggt gtcgttaggc	240
cttcttggaa atctgggtca	ggccctcgaa ctgtgtctcg	ggggcgctga tgctggtcag	300
ctgcacgggc acgttgttct	cgttcaggaa gatgatcagc	ttcttggcggt gtcgtactc	360
ctcgccggcg tggcgaaca	ggaacaggcc ggcccgtcc	aggctgtggg ttagcacca	420
gtgtgtcatg ctcatgtaca	ggttgtctgt ctgcatactcc	ttgttgcacct gtcgttcag	480
cagcttgcgtat	atgtcgccctc cggacttgc	ctcatttctg ttcagettgc	540
gtacttaggg tagtcgttagg	tgccgttctt cacgctgttc	atacactcat cgtcgcactt	600
gtggtagaaac tcgaaggcgc	cattgcccag ttcttcaca	ttgtctctca gtcgtatccg	660
cactttgtcg tacaggatct	tcacattgtct gtcgttggaa	tccagggttc tctcatttcc	720
catcagcacc agcagctcag	cattgtatgt ctgcagctgt	ttgattcccc acactgtcag	780
attcagcaga tgctgtcgag	cttcaatagc tctcagcaga	ttgttctgt gtcgcacaat	840
tccagatcct ccgcctgttc	cattttctcg ctgatttctgg	ctttcctcga tcaggctgt	900
gtatgtgtcg gtgttagttgt	tgtctctct gtcgttgcatt	gtccatttcg gatctccgc	960
gtacacgcta ttcactttgt	tggtgatgcc gtcgaaggct	ttctgtgttag attcttgc	1020
ggcggcatat ccagatccct	gatcattaga gtgggtgttag	ccgttaccacc catcaatcat	1080
tccctgcat ccgccttcaa	taaagccago aatggctcca	aacaggccctc tgctctcaat	1140
ctgaggcaca tttctcagtc	ctgttagccag caccagtct	gatccaatat cttggcgctg	1200
tgtcactgtc acatttcttt	ccaggatgtt atccacttc	tcgggtctat tggcgctg	1260
gtagccaata cagatctgat	cgcctctgac ggctgtaaac	agcagaatca ggtagatgat	1320
ggccat			1326

<210> SEQ ID NO 109

<211> LENGTH: 1362

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 109

atgaagacca tcatgcctt	gagctacatc ttctgcctgg	ccctggccca ggacctgccc	60
ggcaacgaca acagcaccgc	caccctgtgc ctggccacc	acgcccgtgcc caacggcacc	120
ctgggtgaaga ccatcaccga	cgaccagatc gaggtgacca	acgcccacca gctggcgctcc	180
ggcctgaagc tggccacccg	catgcggAAC gtgcccggaa	agcagaccccg gggctgttc	240
ggcgccatcg ccggcttcat	cgagaacggc tgggaggcc	tgtatgcacgg ctggtacggc	300
ttccggcacc agaacagcga	gggcacccggc caggccggcc	acctgaagag caccaggcc	360
gccatcgacc agatcaacgg	caagctgAAC cgggtgatcg	agaagacccgg cggcgatccc	420
gagtgggacc gggagatcaa	caactacacc agcatcatct	acagcctgtat cgaggagac	480
cagaaccacgc	aggagaacgg caccggcgcc	ggcagcggca tcgtcagca	540
ctgctgcggg	ccatcgaggc ccagcagcac	ctgctgcagca tgaccgtgt	600

-continued

```

cagctgcaga gctacaacgc cgagctgctg gtggccctgg agaaccagca caccatcgac      660
ctgaccgaca gcgagatgaa caagctgttc gagaagaccc ggccggcagct gcggggagaac      720
gccgaggaca tgggcaacgg ctgcttcaag atctaccaca agtgcgacaa cgcctgcac          780
gagagcatcc ggaacggcac ctacgaccac gacgtgtacc gggacgaggc cctgaacaac      840
cggttccaga tcaaggggctc cggaggcgac atcatcaagc tgctgaacga gcaggtgaac      900
aaggagatgc agagcagcaa cctgtacatg agcatgacaa gctggtgcta caccacacg      960
ctggacggcg ccggcctgtt cctgttcgac cacggccgg aggagtacga gcacgccaag      1020
aagctgatca tcttcctgaa cgagaacaac gtgcccgtgc agctgaccag catcagcgcc      1080
cccgagcaca agttcgaggg cctgacccag atcttcaga aggctacga gcacgagcag      1140
cacatcagcg agagcatcaa caacatcgtg gaccacgcca tcaagagcaa ggaccacgccc     1200
accttcaact tcctgcagtg gtacgtggcc gagcagcacg aggaggaggt gctgttcaag      1260
gacatcctgg acaagatcga gctgatcggc aacgagaacc acggcctgta cctggccgac      1320
cagtagtga agggcatcgc caagagcagg aagagcggat cc                                1362

```

<210> SEQ ID NO 110

<211> LENGTH: 454

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 110

Met	Lys	Thr	Ile	Ile	Ala	Leu	Ser	Tyr	Ile	Phe	Cys	Leu	Ala	Leu	Gly
1				5				10				15			

Gln	Asp	Leu	Pro	Gly	Asn	Asp	Asn	Ser	Thr	Ala	Thr	Leu	Cys	Leu	Gly
				20				25				30			

His	His	Ala	Val	Pro	Asn	Gly	Thr	Leu	Val	Lys	Thr	Ile	Thr	Asp	Asp
				35				40			45				

Gln	Ile	Glu	Val	Thr	Asn	Ala	Thr	Glu	Leu	Gly	Ser	Gly	Leu	Lys	Leu
				50				55			60				

Ala	Thr	Gly	Met	Arg	Asn	Val	Pro	Glu	Lys	Gln	Thr	Arg	Gly	Leu	Phe
			65			70		75			80				

Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Asn	Gly	Trp	Glu	Gly	Met	Ile	Asp
				85			90			95					

Gly	Trp	Tyr	Gly	Phe	Arg	His	Gln	Asn	Ser	Glu	Gly	Thr	Gly	Gln	Ala
				100			105			110					

Ala	Asp	Leu	Lys	Ser	Thr	Gln	Ala	Ala	Ile	Asp	Gln	Ile	Asn	Gly	Lys
			115			120			125						

Leu	Asn	Arg	Val	Ile	Glu	Lys	Thr	Gly	Gly	Asp	Pro	Glu	Trp	Asp	Arg
			130			135			140						

Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile	Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser
			145			150			155			160			

Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr	Gly	Gly	Ser	Gly	Ile	Val	Gln	
			165			170			175						

Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu
			180			185			190						

Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ser	Tyr	Asn	Ala	Glu
			195			200			205						

Leu	Leu	Val	Ala	Leu	Glu	Asn	Gln	His	Thr	Ile	Asp	Leu	Thr	Asp	Ser
			210			215			220						

-continued

Glu Met Asn Lys Leu Phe Glu Lys Thr Arg Arg Gln Leu Arg Glu Asn
 225 230 235 240
 Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp
 245 250 255
 Asn Ala Cys Ile Glu Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val
 260 265 270
 Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Ser Gly
 275 280 285
 Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln
 290 295 300
 Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His Ser
 305 310 315 320
 Leu Asp Gly Ala Gly Leu Phe Leu Asp His Ala Ala Glu Glu Tyr
 325 330 335
 Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro
 340 345 350
 Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu
 355 360 365
 Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser Glu
 370 375 380
 Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala
 385 390 395 400
 Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu
 405 410 415
 Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu
 420 425 430
 Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys
 435 440 445
 Ser Arg Lys Ser Gly Ser
 450

<210> SEQ ID NO 111
 <211> LENGTH: 1362
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

 <400> SEQUENCE: 111

ggatccgctc ttccctgtct	tggcgatgcc cttcacgtac	tggtcggcca ggtacaggcc	60
gtgggttctcg ttgccgatca	gctcgatctt gtccaggatg	tccttgaaca gcacctcctc	120
ctcgtgtgc tcggccacgt	accactgcag gaagttgaag	gtggcgttgt cttgtcttt	180
gtatggcgtgg tccacgtgt	tgttgatgt ctcgctgtat	tgctgtctgt gctcgtaggc	240
cttctggaaat atctgggtca	ggccctcgaa cttgtgtcg	ggggcgctga tgctggtag	300
ctgcacgggc acgttgttct	cgttcaggaa gatgatcagc	ttcttggcgt gctcgta	360
ctcggggcg tggtcgaaca	ggaacaggcc ggccgcgtcc	aggctgtggg tgtacacca	420
gtgtgcatacg ctcatgtaca	ggttgctgt ctgcata	ttgttcaccc gctcgtag	480
cagcttgcgt	atgtgcgcctc cggagccctt	gatctggAAC cggttgttca	540
ccggatcac	tcgtggcg	ccggatgcctc tcgatgcagg	600
cttgcgttag	atcttgaagc agccgttgc	catgtccctcg gcgttccc	660
ccgggttctc	tcgaacagct tggtcatctc	gctgtcggtc aggtcgatgg	720

-continued

```
<210> SEQ_ID NO 112
<211> LENGTH: 1341
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
```

<400>	SEQUENCE:	112				
atggagaaga	tcgtgtgtct	gctggccatc	gtgagcctgg	tgaagagcga	ccagatctgc	60
atcggttacc	acgccaacaa	cagcacccag	cagggtggaca	ccatcatgga	gaagaacgtg	120
accgtgaccc	acgcccagga	catcgctcc	ggcctgggtc	tggccacccgg	cctgcggaaac	180
agccccccagc	gggagagccg	gcggaagaag	cggggcctgt	tggcgcctat	cggccggcttc	240
atcgaggggc	gtggcaggg	catggtgac	ggctggta	gctaccacca	cagcaacgag	300
cagggcagcg	gtcacgcccgc	cgacaaggag	agcacccaga	aggccatcga	cggcgtgacc	360
aacaagggtga	acagcatcat	cgacaagatg	ggggggcgtac	cegagttggga	ccgggagatc	420
aacaactaca	ccagcatcat	ctacagcctg	atcgaggaga	gccagaacca	gcaggagaac	480
ggcacccggcg	cgggcagcgg	catgtgcag	cagcagaaca	acctgtcg	ggccatcgg	540
gcccagcgc	acctgtgc	gctgaccgt	tggggcatca	agcagctgca	gaccta	600
gcccagatgc	tggtgtgtat	ggagaacag	cgacccctgg	acttccacga	cagcaacgtg	660
aagaacctgt	acgacaagg	gcccgtgcag	ctgcgggaca	acgccaagg	gctggcaac	720
ggctgttgc	agttctacca	caagtgcac	aacgagtgc	tggagagcat	ccggAACGGC	780
acctaact	accccccag	cagcgaggag	gcccggctg	agcggggagga	gatcagctcc	840
ggaggcgaca	tcatcaagct	gctgaacag	caggtaaca	aggagatgc	gagcagcaac	900
ctgtacatga	gcatgagcag	ctgggtctac	acccacacgc	tggacggcgc	cgccctgttc	960
ctgttcgacc	acgcccggcga	ggagtagcag	cacgccaaga	agctgatcat	cttcctgaac	1020
gagaacaac	tgcccggtc	gctgaccgc	atcagcgcc	ccgagcacaa	gttcgagggc	1080
ctgacccaga	tcttccagaa	ggcctacag	cacgaggcgc	acatcagcga	gagcatcaac	1140
aatatcggt	accacgcac	caagagcag	gaccacgc	cattcaactt	cctgcagtgg	1200
tacgtggccg	agcagcacga	ggaggaggt	ctgttcaagg	acatcctg	caagatcgag	1260
ctgatcgcc	acgagaacca	cgccctgtac	ctggccgacc	agtacgtgaa	gggcatacgcc	1320
aagagcgagg	agagcgccat	c				1341

-continued

```

<210> SEQ_ID NO 113
<211> LENGTH: 447
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 113

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1           5          10          15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20          25          30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35          40          45

Gly Ser Gly Leu Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg
50          55          60

Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
65          70          75          80

Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His
85          90          95

His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr
100         105         110

Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp
115         120         125

Lys Met Gly Gly Asp Pro Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
130         135         140

Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Asn
145         150         155         160

Gly Thr Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu
165         170         175

Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly
180         185         190

Ile Lys Gln Leu Gln Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu
195         200         205

Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr
210         215         220

Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn
225         230         235         240

Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser
245         250         255

Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Ala Arg
260         265         270

Leu Lys Arg Glu Glu Ile Ser Ser Gly Gly Asp Ile Ile Lys Leu Leu
275         280         285

Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser
290         295         300

Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe
305         310         315         320

Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile
325         330         335

Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser
340         345         350

Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala
355         360         365

Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp

```

-continued

370 375 380

His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp
 385 390 395 400

Tyr Val Ala Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu
 405 410 415

Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala
 420 425 430

Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
 435 440 445

<210> SEQ ID NO 114

<211> LENGTH: 1341

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 114

```

ggatccgctc ttccctgctct tggcgatgcc cttcacgtac tggtcggcca ggtacaggcc 60
gtgggttcgtc tgccgcgtca gctcgatctt gtccaggatg tccttgaaca gcacctcc 120
ctcgtgtcgc tcggccacgt accactgcag gaagttgaag gtggcgttgt cttgtcttt 180
gtatggcgtgg tccacgtatgt tgttgtatgt ctgcgtgtat tgctgtctgt gctcgtaggc 240
cttcttggaaat atctgggtca ggccctcgaa cttgtgtcg ggggcgtgtga tgctggtc 300
ctgcacgggc acgttgttct cggttgcggaa gatgtatcgc ttcttggcgt gctcgactc 360
ctcggcggcg tggtcgaaca ggaacaggcc ggccgcgtcc aggctgtggg tggtagcacca 420
gctgctcatg ctcatgtaca ggttgtgtct ctgcatactcc ttgttccacct gctcgttcag 480
cagcttgcgtatgtcgtcgccctc cggagctgtat ctcccccgc ttcaaggccggg cctccctcgct 540
gtactggggtagttgttagg tgccgttccg gatgtctcc atgcactcgt tgctgcactt 600
gtggtagaaatcgaaggcgc cggttgcggc ctcccttggc ttgtcccgca gctgcaggcc 660
caccttgcgtacaggttct tcacgttgct gtcgtggaaat tccagggtcc gctcgttctc 720
catcagcacc agcagctcggtcggttgcgtgc ttgtatgcggcc acacgggtcag 780
ctgcagcagg tgctgtgtggg cctcgatggc ccgcagcagg ttgttctgt gctgcacgt 840
ggcgctgcgc cccgcgggtgc cggttctgg ctctccgtca tcaggctgt 900
gatgtatgtgtgtatctcccg gtcccaactcg ggatcgccgc ccatcttgc 960
gatgtatgtgttcacccgttgcgtggcc gtcgtatggcc ttctgggtgc tctccctgtc 1020
ggcgccgttag ccgtgtccct gtcgtgtgtat gtgggtgttag ccgttaccaggc cgtccaccat 1080
ggccctgcgc cccgcgtca tgaaggccgc gatggcgcggc aacaggcccc gcttcttccg 1140
ccggctctcc cgctgggggc tggtccgcag gccgggtggcc agcaccaggc cggagccgt 1200
gtccctggcgtgggtcgtacgg tcacgttctt ctccatgtatgt gtcgtccacct gtcgtgtgt 1260
gttggcgtggcgtggtagccgt tgccagatctg gtcgtcttc accaggctca cgatggccag 1320
cagcagcaccatcttctccat 1341

```

<210> SEQ ID NO 115

<211> LENGTH: 1332

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 115

-continued

atgaaaagtga agctgctggc gctgctgtgt acctttaccc ccacacctacgc cgataaccatc 60
 tggatcggtt accacgccaa caaatgcacc gacaccgtgg ataccgtgt ggaaaagaac 120
 gtgaccgtga cccacagcgt gaacctggga tcaggactga gaatgggtgac cggcctgagg 180
 aatatccccca gcatccagag cagaggcctg tttggcgcca ttgccggctt tatcgaggc 240
 gggatggacag gcatggtgga tgggtggta ggctaccacc accagaatga gcagggatct 300
 ggctatgcgc ccgatcagaa gggccatca aacgcatcac caacaactg 360
 aacagcgtga tcgagaagat gggccggat cctgaatggg acagagagat caacaactac 420
 accagcatca tctacagcct gatcgaggaa agccagaacc agcaggaaaa cggcacaggc 480
 ggccggatctg gaattgtgca gcagcagaac aacctgtga gagccattga ggcccagcag 540
 catctgtgc agctgacagt gtggggcato aacgagctgc agacctacaa cgccgaactc 600
 ctggtcctcc tggaaaatga gaggaccctg gacttccacg acagcaacgt gaagaacctg 660
 tacgagaaag tgaagagcca gctgaagaac aacgccaaag agatggccaa cggctgttc 720
 gagttctacc acaagtgc aaacgagtgc atggaaagcg tgaagaacgg cacctacgac 780
 taccccaagt acagegagga aagcaagctg aaccgggaga agatcgatc cggaggcgac 840
 atcatcaagc tgctgaacga gcaggtgaac aaggagatgc agagcagcaa octgtacatg 900
 agcatgagca gctgggtgcta caccacacgc ctggacggcg cccggctgtt cctgttcgac 960
 cacgcegccc aggagtagcga gcacgccaag aagctgtacaa tcttcctgaa cgagaacaac 1020
 gtgccegtgc agctgaccag catagcgcc cccgagcaca agttcgaggg cctgaccag 1080
 atcttcaga aggccctacga gcacgagcag cacatcagcg agagcatcaa caacatcgt 1140
 gaccacgcca tcaagagcaa ggaccacgccc accttcaact tcttcgtgt gtacgtggcc 1200
 gagcagcacg aggaggaggt gctgttcaag gacatcctgg acaagatcga gctgatcg 1260
 aacgagaacc acggccctgta cctggccgac cagtagtga agggcatcgc caagagcagg 1320
 aagagcggat cc 1332

<210> SEQ ID NO 116
 <211> LENGTH: 444
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 116

Met	Lys	Val	Lys	Leu	Leu	Val	Leu	Leu	Cys	Thr	Phe	Thr	Ala	Thr	Tyr
1															15

Ala	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
20															30

Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
35															45

Leu	Gly	Ser	Gly	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser
50															60

Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly
65															80

Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn
85															95

Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	Asn	Ala
100															110

Ile	Asn	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Gly
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

-continued

115	120	125
Gly Asp Pro Glu Trp Asp Arg	Glu Ile Asn Asn Tyr Thr Ser Ile Ile	
130	135	140
Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln	Glu Asn Gly Thr Gly	
145	150	155
160		
Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu	Arg Ala Ile	
165	170	175
Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp	Gly Ile Lys Gln	
180	185	190
Leu Gln Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu	Glu Asn Glu Arg	
195	200	205
Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu	Tyr Glu Lys Val	
210	215	220
Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly	Asn Gly Cys Phe	
225	230	235
240		
Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu	Ser Val Lys Asn	
245	250	255
Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser	Lys Leu Asn Arg	
260	265	270
Glu Lys Ile Asp Ser Gly Asp Ile Ile Lys Leu Leu	Asn Glu Gln	
275	280	285
Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr	Met Ser Met Ser Ser	
290	295	300
Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly	Leu Phe Leu Phe Asp	
305	310	315
320		
His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu	Ile Ile Phe Leu	
325	330	335
Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile	Ser Ala Pro Glu	
340	345	350
His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys	Ala Tyr Glu His	
355	360	365
Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val	Asp His Ala Ile	
370	375	380
Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln	Trp Tyr Val Ala	
385	390	395
400		
Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu	Asp Lys Ile	
405	410	415
Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu	Ala Asp Gln Tyr	
420	425	430
Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser		
435	440	

```

<210> SEQ ID NO 117
<211> LENGTH: 1332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 117

ggatccgctc ttccctgctct tggcgatgcc cttcacgtac tggtcggcca ggtacaggcc      60
gtgggttctcg ttgcgcgatca gctcgatctt gtccaggatg tccttgaaca gcacccctc      120
ctcgtgtgc tcggccacgt accactgcag gaagttgaag gtggcggtgt ccttgcttt      180
gatggcgatgg tccacgatgt tgttgatgtc ctcgctgatg tgctgctcgt gctcgtaggc      240

```

-continued

cttctggaaatctgggtca ggccctcgaa cttgtgctcg ggggcgctga tgctggtcag	300
ctgcacgggc acgttgttct cggtcaggaa gatgatcage ttcttggcgt gctcgactc	360
ctcggggcg tggtcgaaca ggaacaggcc ggcccgtcc aggctgtggg tgttagcacca	420
gtctgtcatg ctcatgtaca gggtgctgt ctgcatactcc ttgttccacct gctcgttcag	480
cagcttgatg atgtcgccctc cggaatcgat ctctccggg ttcaagtttc tttccctcgct	540
gtacttgggg tagtctgttagg tgccgttctt caccgtttcc atgcactcgt cggttgcatt	600
gtggtagaac tcgaaggcage cggttgcgtat ctctttggcg ttgttcttca gctggcttctt	660
cacttttcg tacagggttct tcacgttctgt gtctggaaat ccagggttcc tctcatttc	720
caggaggacc aggagttcgg cggtttaggt ctgcagctgc ttgtatgcacc acactgtcag	780
ctgcagcaga tgctgttggg cctcaatggc ttcagcagggt ttgttctgt gctgcacaaat	840
tccagatccg ccgcctgtgc cggtttctgt ctgggttctgg ctttccctcga tcaggctgtta	900
gatgtatgtt gtgttagttgt tgatctctct gtccatttca ggatcgccgc ccatttttc	960
gatcacgttgc ttcaattttgt tggtgatgtcc gttgtatggcg ttctgggtgc tttttgtatc	1020
ggcgccatag ccagatccct gtcatttctgt gtgggtgttag ccgttaccacc catccaccat	1080
gcctgtccat ccgcctcgtat taaagccggc aatggcccca aacaggccctc tgctgttggat	1140
gtctggggata ttccctcaggc cggttccat ttcagttctgt gatcccagggt tcacgttgc	1200
ggtcaeggtt acgttctttt ccagcacggt atccacgggt tgggtgttat tggtgggttgc	1260
gttagccgata cagatggat cggcgtaggt ggcggtaaaat gtacacagca gcaccagcag	1320
cttcactttc at	1332

<210> SEQ ID NO 118

<211> LENGTH: 1362

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 118

atgaaaacca tcattgcctt gagctacatc ctgtgcctgg tggtcacaca gaagctggcc	60
ggcaacgata atagcaccgc cacactgtgt ctgggacacc acggccgtgcc taatggcacc	120
atcgtaaaa caatcaccaa cgaccagatc gaagtgtacca atgcccacaga gctgggttcc	180
ggcctgaagc tggccacccgg catgagaaat gtggccgaga agcagaccag aggcatctt	240
ggcgccatttgc cggttttat cgagaatggc tgggaggggaa tggtggatgg gtggtaaggc	300
ttcagacacc agaatagcga gggatttggc caggccggcc atctgaaatc taccaggcc	360
gccatcgacc agatcaacgg caagctgaac aggctgtatcg gcaagaccgg cggcgatccc	420
gagttggacc gggagatcaa caactacacc agcatcatct acagcctgtat cgaggagac	480
cagaaccaggc aggagaacgg caccggccggc ggcagcggca tcgtgcagca gcagaacaac	540
ctgctgcggg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgt gggcatcaag	600
cagctgcaga gtcataatgc cgaactgtgt gtgccttggg aaaaccagca cacaatttgc	660
ctgacagaca gtgagatgaa taagctgttc gagaaaaccat agaaggcagct gagagaaaaac	720
ggcgaggaca tggcaacgg ctgcttcaag atctaccaca agtgcgacaa cgcctgcac	780
ggcagcatca gaaacggcac ctacgaccac gacgtgtaca gagatggggc cctgaacaac	840
cggtttcaga tcaagggctc cggaggcgcac atcatcaagc tgctgaacga gcaggtgaac	900

-continued

aaggagatgc agagcagcaa cctgtacatg	960
agcatgagca gctggtgcta cacccacagc	
ctggacggcg ccggcctgtt cctgttcgac	1020
cacgcccggc aggagtacga gcacgccaag	
aagctgatca tcttcctgaa cgagaacaac gtgcccgtgc	1080
agctgaccag catcagcgcc	
cccgagcaca agttcgaggg cctgacccag atcttccaga	1140
aggcctacga gcacgagcag	
cacatcagcg agagcatcaa caacatcgtg gaccacgcca	1200
tcaagagcaa ggaccacgcc	
accttcaact tcctgcagtg gtacgtggcc gagcagcagc	1260
aggaggaggt gctgttcaag	
gacatcctgg acaagatcga gctgateggc aacgagaacc acggcctgta	1320
cctggccgac	
cagtagtgc agggcattgc caagagcagg aagagcggat cc	1362

<210> SEQ ID NO 119

<211> LENGTH: 454

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 119

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Thr			
1	5	10	15

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly			
20	25	30	

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp			
35	40	45	

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Gly Ser Gly Leu Lys Leu			
50	55	60	

Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe			
65	70	75	80

Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp			
85	90	95	

Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Ile Gly Gln Ala			
100	105	110	

Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys			
115	120	125	

Leu Asn Arg Leu Ile Gly Lys Thr Gly Gly Asp Pro Glu Trp Asp Arg			
130	135	140	

Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser			
145	150	155	160

Gln Asn Gln Gln Glu Asn Gly Thr Gly Gly Ser Gly Ile Val Gln			
165	170	175	

Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu			
180	185	190	

Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ser Tyr Asn Ala Glu			
195	200	205	

Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp Ser			
210	215	220	

Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn			
225	230	235	240

Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp			
245	250	255	

Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val			
260	265	270	

Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Ser Gly			
275	280	285	

-continued

Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln
 290 295 300

 Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His Ser
 305 310 315 320

 Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu Tyr
 325 330 335

 Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro
 340 345 350

 Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu
 355 360 365

 Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser Glu
 370 375 380

 Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala
 385 390 395 400

 Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu
 405 410 415

 Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu
 420 425 430

 Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys
 435 440 445

 Ser Arg Lys Ser Gly Ser
 450

-continued

cattccctcc cagccattct cgataaagcc ggcaatggcg ccaaagatgc ctctggctcg	1140
cttctcgggc acatttctca tgccgggtggc cagttcagg ccggageccca gctctgtggc	1200
atgggtcact tcgatctggt cgttgttgat tgttttcacg atgggtccat taggcacggc	1260
gtggtgtccc agacacagtg tggcggtgtc attatcggtt ccgggcagct tctgtgtgaa	1320
caccadggcac aggatgttaqc tcaggccaat qatggtttc at	1362

```
<210> SEQ ID NO 121
<211> LENGTH: 1362
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 121

atgaaaaacc taattgcgt gtcctacata ctgtgtctgg tggggccca gaaactgccc 60
ggcaatgaca actcaacagc cacgctctgc ttggggcacc atgcccgtccc taacgggacc 120
attgtaaaa ccattactaa cgatcagata gaggtgacta atgcccacccg gctgggttcc 180
ggcttgaac tatgcgcaat gtccccgaaa aacagacccg cgggatattt 240
ggggctatcg caggctttat cgagaatggc tgggaaggga tggtgatgg ttggatgg 300
tttagacatc aaaactccga aggccagggc caggctgccc atctcaagag cacgcaggcc 360
gctatagatc agatcaatgg aaagctcaac agactgatcg ggaaaacccg cggcgatccc 420
gagtgggacc gggagatcaa caactacacc agcatcatcg acagcctgtat cgaggagagc 480
cagaaccaggc aggagaacgg cacccggcggc ggcagcggca tcgtgcagca gcagaacaac 540
ctgctggggc ccatcgaggg ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 600
cagctgcagt cctacaatgc cgagctgctg gtggctctgg agaatcagca cactattgac 660
ctgaccgatt cagagatgaa caaactttt gagaagacga agaaggagct tagagaaaat 720
gcagaggaca tggggAACGG atgcttaaa atatatcata agtgtgataa tgcctgcattc 780
ggatcaatta gaaatggtac ctatgatcac gatgttaca gggacgaaac gctgaataac 840
aggttccaga taaaaggctc cggaggcgac atcatcaagc tgctgaacga gcagggtgaac 900
aaggagatgc agagcagcaa cctgtacatcg agcatgagca gctgggtcta caccacagc 960
ctggacggcg cccgcctgtt cctgttcgac cacgcccgg aggagtagcga gcacgcctaa 1020
aagctgatca tcttcctgaa cgagaacaac gtgcccgtgc agctgaccag catcagcgcc 1080
cccgagcaca agttcgaggg cctgaccagg atcttccaga aggccctacga gcacgcggc 1140
cacatcagcg agagcatcaa caacatcgatg gaccacgcca tcaagagcaa ggaccacgccc 1200
accttcaact tcctgcagtg gtacgtggcc gagcagcaccg aggaggaggt gctgttcaag 1260
gacatcctgg acaagatcga gctgatcgcc aacgagaacc acggcctgtta cctggccgac 1320
cagtagtgcg aagggcatcgcc caagagcagg aagagccgtt cc 1362

```
<210> SEQ ID NO 122
<211> LENGTH: 454
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 122

Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Leu Cys Leu Val Phe Ala
1 5 10 15

-continued

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
20 25 30

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
35 40 45

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Gly Ser Gly Leu Lys Leu
50 55 60

Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe
65 70 75 80

Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp
85 90 95

Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala
100 105 110

Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys
115 120 125

Leu Asn Arg Leu Ile Gly Lys Thr Gly Gly Asp Pro Glu Trp Asp Arg
130 135 140

Glu Ile Asn Asn Tyr Thr Ser Ile Ile Tyr Ser Leu Ile Glu Glu Ser
145 150 155 160

Gln Asn Gln Glu Asn Gly Thr Gly Gly Ser Gly Ile Val Gln
165 170 175

Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu
180 185 190

Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ser Tyr Asn Ala Glu
195 200 205

Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp Ser
210 215 220

Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn
225 230 235 240

Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp
245 250 255

Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val
260 265 270

Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Ser Gly
275 280 285

Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln
290 295 300

Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His Ser
305 310 315 320

Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu Tyr
325 330 335

Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro
340 345 350

Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu
355 360 365

Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser Glu
370 375 380

Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala
385 390 395 400

Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu
405 410 415

Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu
420 425 430

-continued

Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys
 435 440 445

Ser Arg Lys Ser Gly Ser
 450

<210> SEQ ID NO 123
 <211> LENGTH: 1362
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 123

ggatccgctc ttccctgtct tggcgatgcc cttcacgtac tggtcggcca ggtacaggcc	60
gtgggttctcg ttggcgatca gctcgatctt gtccaggatg tccttgaaca gcacctctc	120
ctcggtgtgc tcggccacgt accactgcag gaagttgaag gtggcggtgg cttgtctt	180
gatggcggtgg tccacgatgt tggtgatgtc tcgcgtatgt tgctgtctgt gtcgttaggc	240
cttctggaaag atctgggtca ggcccctcgaa cttgtgtcg gggcgctga tgctggtcag	300
ctgcacgggc acgttgttct cggtcaggaa gatgatcagc ttcttggcgt gtcgtactc	360
ctcggcgccg tggcgaaaca ggaacaggcc ggccgcgtcc aggctgtggg tgttagcacca	420
gctgctcatg ctcatgtaca gggtgctgtc tcgcatttc ttgttcaccc gtcgttcag	480
cagcttgatg atgtcgccctc cggagccctt tatctggAAC ctgttattca gcgctcg	540
cctgtaaaca tcgtgtatcat aggtaccatt tctaattgtat ccgtatgcagg cattatcaca	600
cttatgtatat attttaaagc atccgttccc catgtccctc gcattttctc taagctgctt	660
cttcgtcttc tcaaaaagtt tggtcatctc tgaatcggtc aggtcaatag tgtgtgtt	720
cgtccagagcc accagcagct cggcattgtaa ggactgcagc tgcttgatgc cccacacgg	780
cagctgcagc aggtgtgtct gggcctcgat ggccgcgc aggttgttct gtcgtgcac	840
gatgcccgtcg cggccgcgg tggcgatctc ctgcgtggtc tggctctcct cgtatcaggct	900
gttagatgtat ctgggtgtat tggtgatctc ccggatccac tggggatcgc cggcggttt	960
cccgatcagt ctgttgagct ttccattgtat ctgtatatac gggccctgcg tgcttttgag	1020
atcggcagcc tggcctctgc ctccggagtt ttgtatgtata aaaccatacc aaccatccac	1080
catcccttcc cagccattct cgataaagcc tgcgtatgc ccaaataatcc cgcgggtctg	1140
ttttcgggg acattgcgc taccggcgc cagttcaag ccggagccca gtcgggtggc	1200
attagtcacc tctatctgtat cgttagtaat ggtttcaca atggtcccgt tagggacggc	1260
atgggtcccc aagcagagcg tggctgtga gttgtcattt ccggcagtt tctggccaaa	1320
caccagacac agtatgttagg acagcgcaat tatggtttc at	1362

<210> SEQ ID NO 124
 <211> LENGTH: 1332
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 124

atgaaggcca tcatcggtct gctgtatggtg gtcacaagca acggccatag aatctgtacc	60
ggcatcaacca gcaagcaatag ccctcacgtc gtggaaaacac ctacacaggcg cgaagtgaat	120
gtgaccggcg tgatccctct gggatcgaga ctgaagctgg ccaatggcac aaagtataga	180
cctccagccca agctgtgaa agagagaggc tttttggag ctatcgccgg ctttctggaa	240

-continued

ggcgatggg	aggaaatgtat	tgctggatgg	catggctaca	catctcatgg	cgcacatggc	300
gtggcagtgg	ctgctgatct	gaaatctaca	caggaaggcca	tcaacaagat	caccaagaac	360
ctgaacagcc	tgagcgagct	ggaaggaggc	gaccccggat	gggatcgca	aatcaacaac	420
tacacatcta	tcatctacag	tctgatttag	gaaagccaga	accagcagga	aatgggact	480
gggggaggct	ccggaatcgt	gcagcagcag	aacaatctgc	tgcgagccat	tgaagctcag	540
cagcacatgc	tgcgactgac	agtgtgggc	atcaagcgc	tgcaaggagg	ccagattgaa	600
ctggctgtgc	tgctgtctaa	cgagggcata	atcaatagcg	aggacgaaca	tctgtggcc	660
ctggaaagaa	agctgaagaa	gatgtggga	cctagcgcgc	tggaaatcgg	caatggatgc	720
ttttagacaa	agcacaagtg	caaccagacc	tgcctggata	gaattgccgc	cggaacattt	780
gatgccggcg	agttttctct	gcccaccttc	gatagcctga	atatcacatc	cggaggcgac	840
atcatcaagc	tgctgaacga	gcaggtgaac	aaggagatgc	agagcagcaa	cctgtacatg	900
agcatgagca	gctgggtgcta	caccacacgc	ctggacggcg	cggccctgtt	cctgttcgac	960
cacgcccggc	aggagtaacga	gcacgccaag	aagctgatca	tcttcctgaa	cgagaacaac	1020
gtgccccgtgc	agctgaccag	catcagcgc	cccgagcaca	agttcgaggg	cctgaccatg	1080
atcttcaga	aggcctacga	gcacgagcag	cacatcagcg	agagcatcaa	caacatcgt	1140
gaccacgcca	tcaagagcaa	ggaccacgccc	accttcaact	tctgcagtg	gtacgtggcc	1200
gagcagcagc	aggaggaggt	gctgttcaag	gacatcctgg	acaagatcga	gctgatcgcc	1260
aacgagaacc	acggcctgta	cctggccgac	cagtaacgt	agggcatcgc	caagagcagg	1320
aagageggat	cc					1332

<210> SEQ ID NO 125

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 125

Met	Lys	Ala	Ile	Ile	Val	Leu	Leu	Met	Val	Val	Thr	Ser	Asn	Ala	Asp
1					5			10			15				

Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys
					20			25			30				

Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Gly
					35			40			45				

Ser	Gly	Leu	Lys	Leu	Ala	Asn	Gly	Thr	Lys	Tyr	Arg	Pro	Pro	Ala	Lys
					50			55			60				

Leu	Leu	Lys	Glu	Arg	Gly	Phe	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Leu	Glu
					65			70			75			80	

Gly	Gly	Trp	Glu	Gly	Met	Ile	Ala	Gly	Trp	His	Gly	Tyr	Thr	Ser	His
					85			90			95				

Gly	Ala	His	Gly	Val	Ala	Val	Ala	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Glu
					100			105			110				

Ala	Ile	Asn	Lys	Ile	Thr	Lys	Asn	Leu	Asn	Ser	Leu	Ser	Glu	Leu	Glu
					115			120			125				

Gly	Gly	Asp	Pro	Glu	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile
					130			135			140				

Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr
					145			150			155			160	

-continued

Gly Gly Gly Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala
165 170 175

Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys
180 185 190

Gln Leu Gln Gly Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu
195 200 205

Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys
210 215 220

Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys
225 230 235 240

Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala
245 250 255

Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser
260 265 270

Leu Asn Ile Thr Ser Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln
275 280 285

Val Asn Lys Glu Met Gln Ser Ser Asn Leu Tyr Met Ser Met Ser Ser
290 295 300

Trp Cys Tyr Thr His Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp
305 310 315 320

His Ala Ala Glu Glu Tyr Glu His Ala Lys Lys Leu Ile Ile Phe Leu
325 330 335

Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu
340 345 350

His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His
355 360 365

Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile
370 375 380

Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala
385 390 395 400

Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile
405 410 415

Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr
420 425 430

Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser
435 440

<210> SEQ ID NO 126
<211> LENGTH: 1332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 126

ggatccgctc ttccctgtct	tggcgatgcc cttcacgtac	tggtcggcca ggtacaggcc	60
gtggttctcg ttgcgcgtca	gctcgatctt gtccaggatg	tccttgaaaca gcacctctc	120
ctcgtgtgc tcggccacgt	accactgcag gaagttgaag	gtggcgtgg ctttgtcttt	180
gatggcgtgg tccacgtatgt	tgttgatgtc	ctcgatgtatgt tgctgtcggt	240
cttctggaag atctgggtca	ggccctcgaa cttgtgtcg	ggggcgctga tgctggtagc	300
ctgcacgggc acgttgttct	cgttcaggaa gatgatcagc	ttcttggcggt gctcgtaactc	360
ctcggccggcg tggtcgaaca	ggaacaggcc ggcccgtcc	aggctgtggg tgttagcacca	420
gctgctcatgt ctcatgtaca	ggttgctgtct	ctgcatactcc ttgttcaccc gctcggttcag	480

291

292

-continued

cagcttgcgtat atgtcgccctc cggatgtgat attcaggccta tgcagggtgg gcagagaaaa	540
ctcgccggca tcaaattgttc cggcggaat tctatccagg caggctctgg tgcacttgg	600
cttttgtctca aagcatccat tgccgatttc cacggcgtta ggtcccaagca tcttttcag	660
ctttcttcc agggccagca gatgttcgtc ctgcgtatttgc atgatgcct cgtagacag	720
cagcacagcc agttcaatct ggctccccctg cagctgttgc atgcccaca ctgtcagctg	780
cagcagggtgc tgctgagctt caatggctcg cagcagatttgc ttctgtgt gcacgattcc	840
ggagccccc ccagtccat tctctgtgt gttctgttttgc tctcaatca gactgttagat	900
gatacatgtt tagttgttga tttegcgtat ccactcgaaa tgcctctt ccagtcgtct	960
caggctgttc aggtttttgg tgcatttttttgc gatggcttc tgtgttagatt tcagatcagc	1020
agccactgcc acgccccatgtg cggccatgaga tgtgttagcca tgccatccag caatattcc	1080
ctccccatccg ccttccagaa agccggcgat agctccaaaa aagccctctt ctgttgcgtat	1140
cttggcttgc ggtctataact ttgtgcatttgc ggcagcttc agtccgtatcc accagggat	1200
cacggccggtc acatttactt cggccatgttgc agctgttttgc acgacgtgag ggctattgt	1260
gctgggtatccat cccgtacaga ttctatccggc gttgtgttgc accaccatca gcagacgtat	1320
qatqqcccttc at	1332

```
<210> SEQ ID NO 127
<211> LENGTH: 1332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 127
atgaaggcca tcategtgtc gctgatggtg gtgaccagca acggccatag aatctgcacc 60
ggcatcacca gcagcaatag cccccatgtg gtgaaaacag ccaccaggcg cgaagtgaat 120
gtgacaggcg tgatccctct gggatcagga ctgaagctgg ccaatggcac caagtacaga 180
cctcccccca agctgctgaa agagagaggc ttctttggcg ccattgcgg atttctggaa 240
ggcggtctgg agggaatgat tgccggctgg cacggctata catctcatgg ggcccatggc 300
gtggctgtgg cgcgcgatct gaagtctacc caggaagcca tcaacaagat caccaagaac 360
ctgaacagcc tgagcgcagct ggaaggaggc gaccccggat gggatcgcga aatcaacaac 420
tacacatcta tcatctacag tctgatttag gaaagccaga accagcagga gaatggact 480
ggggggaggct cgggaatcgt gcagcagcag aacaatctgc tgcgagccat tgaagctcag 540
cagcacctgc tgcagctgac agtgtgggc atcaagcagc tgcaggggat ccagattgaa 600
ctggccgtgc tgctgtccaa cgaggccatc atcaacagcg aggtatgaaca cctgtggcc 660
ctggaaacgga agctgaagaa gatgctggc cttctgcgg tggagatcgg caacggctgc 720
ttcgagacaa agcacaagtgc caaccagacc tgcctggata gaatgcggc tggcaccttc 780
aatgcggcgc agttcagcct gcctaccttc gacagcctga atatcaccttc cggaggccgac 840
atcatcaagc tgctgaacgc gcagggtgaac aaggagatgc agagcagcaa cctgtacatg 900
agcatgagca gctgggtcta caccacagc ctggacggcg cggccgttt cctgttcgac 960
cacgcccggcaggaggtacga gcacgcacaa aagctgatca tttctctgaa cgagaacaac 1020
gtgcccgtgc agctgaccag catcagcgc cccgagcaca agttcgaggg cctgaccag 1080
atcttccaga aggctacga gcacgcacaa aacatcagcg agagcataa caacatcgtg 1140

-continued

gaccacgcca tcaagagcaa ggaccacgcc accttcaact tcctgcagt gtacgtggcc 1200
 gagcagcagc aggaggagggt gctgttcaag gacatcctgg acaagatcga gctgatccgc 1260
 aacgagaacc acggcctgta cctggccgac cagtacgtga agggcatcgc caagagcagg 1320
 aagagcggat cc 1332

<210> SEQ ID NO 128
 <211> LENGTH: 444
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic
 <400> SEQUENCE: 128

Met	Lys	Ala	Ile	Ile	Val	Leu	Leu	Met	Val	Val	Thr	Ser	Asn	Ala	Asp
1															
															15
Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys
															30
Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Gly
															45
Ser	Gly	Leu	Lys	Leu	Ala	Asn	Gly	Thr	Lys	Tyr	Arg	Pro	Pro	Ala	Lys
															60
Leu	Leu	Lys	Glu	Arg	Gly	Phe	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Leu	Glu
															80
Gly	Gly	Trp	Glu	Gly	Met	Ile	Ala	Gly	Trp	His	Gly	Tyr	Thr	Ser	His
															95
Gly	Ala	His	Gly	Val	Ala	Val	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Glu	
															110
Ala	Ile	Asn	Lys	Ile	Thr	Lys	Asn	Leu	Asn	Ser	Leu	Ser	Glu	Leu	Glu
															125
Gly	Gly	Asp	Pro	Glu	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Ile
															130
Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Asn	Gly	Thr
															145
															150
Gly	Gly	Gly	Ser	Gly	Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala
															165
															170
Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys
															180
Gln	Leu	Gln	Gly	Ser	Gln	Ile	Glu	Leu	Ala	Val	Leu	Leu	Ser	Asn	Glu
															195
Gly	Ile	Ile	Asn	Ser	Glu	Asp	Glu	His	Leu	Leu	Ala	Leu	Glu	Arg	Lys
															210
Leu	Lys	Lys	Met	Leu	Gly	Pro	Ser	Ala	Val	Glu	Ile	Gly	Asn	Gly	Cys
															225
Phe	Glu	Thr	Lys	His	Lys	Cys	Asn	Gln	Thr	Cys	Leu	Asp	Arg	Ile	Ala
															245
Ala	Gly	Thr	Phe	Asn	Ala	Gly	Glu	Phe	Ser	Leu	Pro	Thr	Phe	Asp	Ser
															260
Leu	Asn	Ile	Thr	Ser	Gly	Gly	Asp	Ile	Ile	Lys	Leu	Leu	Asn	Glu	Gln
															275
Val	Asn	Lys	Glu	Met	Gln	Ser	Ser	Asn	Leu	Tyr	Met	Ser	Met	Ser	Ser
															290
Trp	Cys	Tyr	Thr	His	Ser	Leu	Asp	Gly	Ala	Gly	Leu	Phe	Leu	Phe	Asp
															305
His	Ala	Ala	Glu	Glu	Tyr	Glu	His	Ala	Lys	Lys	Leu	Ile	Ile	Phe	Leu

-continued

325	330	335	
Asn Glu Asn Asn Val Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu			
340	345	350	
His Lys Phe Glu Gly Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His			
355	360	365	
Glu Gln His Ile Ser Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile			
370	375	380	
Lys Ser Lys Asp His Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala			
385	390	395	400
Glu Gln His Glu Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile			
405	410	415	
Glu Leu Ile Gly Asn Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr			
420	425	430	
Val Lys Gly Ile Ala Lys Ser Arg Lys Ser Gly Ser			
435	440		

<210> SEQ ID NO 129

<211> LENGTH: 1332

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 129

ggatccgctc ttccctgtct tgccgatgcc cttcacgtac tggcgccca ggtacaggcc	60
gtgggttcgc ttgcgcgatca gctcgatctt gtccaggatg tccttgaaca gcacctcc	120
ctcggtgtgc tcggccacgt accactgcag gaagttgaag gtggcggtgt cttgtctt	180
gatggcggtgg tccacgatgt tggtgatgt ctgcgtatg tgctgtctgt gtcgttaggc	240
cttcttggaa atctgggtca ggccctcgaa cttgtgtcg gggcgctga tgctggtag	300
ctgcacgggc acgttgttct cggtcaggaa gatgtacgtc ttcttggcggt gtcgtactc	360
ctcggcgccg tggcgaaaca ggaacaggcc gggccgtcc aggctgtggg tggcgaccca	420
gtgtgtcatcg ctcatgtaca gggtgtgt ctgcgtatcc ttgttacact gtcgtttag	480
cagctttagt atgtcgccctc cggaggtat attcaggctg tgcggatgt gcaggctgaa	540
ctcgccggca ttgaagggtgc cagggcgat tctatccagg caggctgtgt tgcaatttg	600
ctttgtctcg aagcagccgt tgccgatctc cacggcagaa gggcccagca tcttttc	660
cttccgttcc agggccagca ggtgttcatc ctgcgtgtt atgtgtccct cggtggacag	720
cagcacggcc agttcaatct gggaccctg cagctgttg atgccccaca ctgtcagtg	780
cagcagggtgc tgctgagtt caatggctcg cagcagatgg ttctgtgtct gcacgatcc	840
ggagcctccc ccagtccat tctccgtcg gtctggctt tcctcaatca gactgttagat	900
gatagatgtg tagttgtta ttccgcgtc ccactcgcccc tggccctccctt ccagtcgt	960
caggctgttc aggttcttgg tggatcttggat gatggcttcc tgggttagact tcagatcc	1020
ggccacagcc acgcccatttggg ccccatgaga tgtatagccg tgccagccgg caatcatcc	1080
ctccccagccg cttcccgaaa atccggcaat ggccggaaag aagccctctt ctgtcagcag	1140
cttggcgccg ggtctgtact tggcgccatt ggccagttcc agtccgtatc ccagaggat	1200
cacgcctgtc acattcaattt cggccctgggt ggctgttttc accacatggg ggctattgt	1260
gctgggtatc ccgggtcaga ttctatcgcc gttgtggtc accaccatca gcagcacat	1320
gatggcatttc at	1332

297

What is claimed is:

1. A nanoparticle comprising a fusion protein, wherein the fusion protein comprises a monomeric ferritin subunit protein joined to an influenza hemagglutinin (HA) protein, such that the nanoparticle comprises influenza virus HA protein trimers on its surface.

2. The nanoparticle of claim **1**, wherein the monomeric ferritin subunit protein is a monomeric subunit of a *Helicobacter pylori* ferritin protein.

3. The nanoparticle of claim **1**, wherein the hemagglutinin protein is from an influenza virus selected from the group consisting of A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), and B/Brisbane/60/2008 (2008 Bris, B).

4. The nanoparticle of claim **1**, wherein the hemagglutinin protein comprises a region selected from the group consisting of:

a) a region corresponding to amino acids 1-519 of SEQ ID NO:8;

b) a region comprising amino acids 1-519 of SEQ ID NO:8;

c) a region comprising an amino acid sequence at least about 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98; and,

d) a region comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

5. The nanoparticle of claim **1**, wherein the nanoparticle elicits an immune response to:

a) an influenza virus strain that is heterologous to the strain of influenza virus from which the hemagglutinin protein was obtained; or,

b) a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

6. The nanoparticle of claim **1**, wherein the fusion protein comprises an amino acid sequence selected from the group consisting of

SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID

298

NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

7. A method to produce a vaccine against influenza virus, the method comprising a) expressing a fusion protein comprising a monomeric ferritin protein joined to an influenza hemagglutinin protein under conditions such that the fusion proteins form a nanoparticle displaying hemagglutinin trimers on its surface, and b) recovering the nanoparticle.

8. A method to vaccinate an individual against influenza, the method comprising administering a vaccine produced according to the method of claim **7** to an individual such that the nanoparticle elicits an immune response against influenza virus.

9. The method of claim **8**, wherein the nanoparticle elicits an immune response to an influenza virus strain that is heterologous to the sub-type of influenza virus from which the hemagglutinin protein was obtained.

10. The method of claim **8**, wherein the nanoparticle elicits an immune response to an influenza virus strain that is heterologous to the strain of influenza virus from which the hemagglutinin protein was obtained.

11. The method of claim **8**, wherein the nanoparticle elicits an immune response to an influenza virus that is antigenically divergent from the influenza virus from which the hemagglutinin protein was obtained.

12. The method of claim **8**, wherein administering comprises administering to the individual a first vaccine composition and then at a later time, administering a second vaccine composition comprising a nanoparticle that comprises an HA-SS-ferritin fusion protein.

13. The method of claim **12**, wherein the HA portion of the HA SS-ferritin fusion protein comprises a region selected from the group consisting of:

a) a region corresponding to amino acids 1-519 of SEQ ID NO:8;

b) a region comprising amino acids 1-519 of SEQ ID NO:8

c) a region comprising an amino acid sequence at least about 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98, wherein the HA SS-ferritin fusion protein elicits an immune response to an influenza virus; and,

d) a region comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

14. The method of claim **12**, wherein the HA SS-ferritin fusion protein comprises an amino acid sequence selected from the group consisting of:

SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID

299

NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

15. The method of claim **12**, wherein the first vaccine composition comprises a nanoparticle comprising a hemagglutinin protein from an influenza virus selected from the group consisting of A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), and B/Brisbane/60/2008 (2008 Bris, B).

16. The method of claim **15**, wherein the hemagglutinin protein comprises a region selected from the group consisting of:

a) a region comprising an amino acid sequence at least about 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98;

b) a region comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98;

c) a region corresponding to amino acids 1-519 of SEQ ID NO:8; and,

d) a region comprising amino acids 1-519 of SEQ ID NO:8.

17. The method of claim **12**, wherein the first vaccine composition comprises an HA-ferritin fusion protein comprising an amino acid sequence selected from the group consisting of

SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113 SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

18. A fusion protein comprising a monomeric ferritin subunit protein joined to an influenza hemagglutinin protein, wherein the monomeric ferritin subunit protein comprises a domain that allows the fusion protein to self-assemble into nanoparticles.

300

19. The fusion protein of claim **18**, wherein the monomeric subunit is a monomeric subunit of a *Helicobacter pylori* ferritin protein.

20. The fusion protein of claim **18**, wherein the hemagglutinin protein comprises a region selected from the group consisting of:

a) a region comprising an amino acid sequence comprising at least 25 amino acids from a second hemagglutinin protein from an influenza virus selected from the group consisting of A/New Caledonia/20/1999 (1999 NC, H1), A/California/04/2009 (2009 CA, H1), A/Singapore/1/1957 (1957 Sing, H2), A/Hong Kong/1/1968 (1968 HK, H3), A/Brisbane/10/2007 (2007 Bris, H3), A/Indonesia/05/2005 (2005 Indo, H5), B/Florida/4/2006 (2006 Flo, B), A/Perth/16/2009 (2009 Per, H3), A/Brisbane/59/2007 (2007 Bris, H1), and B/Brisbane/60/2008 (2008 Bris, B);

b) a region corresponding to amino acids 1-519 of SEQ ID NO:8;

c) a region comprising amino acids 1-519 of SEQ ID NO:8;

d) a region comprising an amino acid sequence at least about 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98; and,

e) a region comprising an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29, SEQ ID NO:32, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:77, SEQ ID NO:80, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:92, SEQ ID NO:95 and SEQ ID NO:98.

21. The fusion protein of claim **18**, wherein the fusion protein comprises an amino acid sequence selected from the group consisting of

SEQ ID NO:41, SEQ ID NO:44, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, SEQ ID NO:59, SEQ ID NO:62, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:101, SEQ ID NO:104 SEQ ID NO:107 SEQ ID NO:110 SEQ ID NO:113, SEQ ID NO:116 SEQ ID NO:119 SEQ ID NO:122 SEQ ID NO:125 and SEQ ID NO:128.

* * * * *